Reconciling land / ocean moisture transport variability in reanalyses with P-ET in observationally-driven land surface models

Franklin R. Robertson[1], Michael G. Bosilovich[2] and Jason B. Roberts[1]

[1]NASA / Marshall Space Flight Center, Earth Science Office
MSFC, AL 35812

[2]GSFC Global Modeling and Assimilation Office
Greenbelt, MD 20771

Corresponding author: Franklin R. Robertson
(pete.robertson@nasa.gov)
tel: (256) 961-7836

Submitted to J. Climate (MERRA-2 Special Collection)

Aug 22, 2016

Abstract

Vertically-integrated atmospheric moisture transport from ocean to land, VMFC, is a dynamic component of the global climate system but remains problematic in atmospheric reanalyses with current estimates having significant multi-decadal global trends differing even in sign. Regional VMFC trends over continents are especially uncertain. Continual evolution of the global observing system, particularly step-wise improvements in satellite observations, has introduced discrete changes in the ability of data assimilation to correct systematic model biases, manifesting as non-physical variability. Land Surface Models (LSMs) forced with observed precipitation, P, and near-surface meteorology and radiation provide estimates of evapotranspiration, ET. Since variability of atmospheric moisture storage is small on interannual and longer time scales, VMFC = P-ET is a good approximation and LSMs can provide an alternative estimate. However, heterogeneous density of rain gauge coverage, especially the sparse coverage over tropical continents, remains a serious concern.

Rotated Principal Component Analysis (RPCA) with pre-filtering of VMFC to isolate the artificial variability is used to investigate artifacts in five reanalysis systems. This procedure, though ad hoc, enables useful VMFC corrections over global land. P-ET estimates from seven different LSMs are evaluated and subsequently used to confirm the efficacy of the RPCA-based adjustments. Global VMFC trends over the period 1979-2012 ranging from 0.07 to -0.03 mmd^{-1} $decade^{-1}$ are reduced by the adjustments to 0.016 mmd^{-1} $decade^{-1}$, much closer to the LSM P-ET estimate (0.007 mmd^{-1} $decade^{-1}$). Neither is significant at the 90 percent level. ENSO-related modulation of VMFC and P-ET remains the largest global interannual signal with mean LSM and adjusted reanalysis time series correlating at 0.86.

1. Introduction

Moisture transport to land from the global oceans is a crucial process linking the global water and energy cycles and is also at the heart of societal concerns regarding terrestrial water availability, food security, exposure to extreme weather events, and climate change. Recent best estimates of the net atmospheric transport of water to land (Rodell et al. 2014) put the climatological amount at 45.8 +/- 6.7 x 10^3 km^3yr^{-1}, or about 40% of precipitation falling over land. The remainder of land precipitation arises from moisture recycling via evapotranspiration, ET (Eltahir and Bras, 1994; Trenberth 1999; Bosilovich and Schubert, 2002). The *variability* of this transport and its potential long-term trend at regional scales are emerging as a prime concern. Tropical circulations linked to sea-surface temperature (SST) variability exert first order controls on the delivery of water to land by virtue of El Niño / Southern Oscillation (ENSO) events (Ropeleski and Halpert, 1987; Dai and Wigley, 2000; Gu et al, 2007; Robertson et al, 2014). Mid-latitude storm track changes embodying teleconnections with tropical forcing also have significant variations at higher latitudes. Over longer time scales Pacific Decadal Variability (PDO / PDV), (e.g. Power et al, 1999; Dai 2013; Lyon et al, 2013) and other basin scale phenomena (e.g. the Atlantic Multi-decadal Oscillation, AMO, Enfield et al, 2001; Sutton and Hodgson, 2005; Ting et al, 2011) also modulate moisture transport. Anthropogenic radiative forcing changes and the consequent hydrologic cycle effects are expected to produce regional variations, encapsulated in the "wet get wetter / dry get drier" paradigm (Chou and Neelin, 2004) wherein hydrologic extremes are expected to increase. As yet, evidence for this behavior in observational data sets is weak at best (Greve et al, 2014). There is also substantial uncertainty as to trends in soil moisture dryness depending on diagnostic approaches and choice of observed precipitation and surface meteorological forcing (Dai, 2011; Sheffield et al, 2012, Trenberth et al, 2014). Untangling the

role of these varied water and energy cycle mechanisms and their relationship to moisture transport continues to be a challenging task.

Moisture transport syntheses are routinely produced by reanalysis efforts (e.g. Kalnay et al, 1996; Onogi et al, 2007; Saha et al, 2010; Dee et al, 2011; Rienecker et al, 2011; Kobayashi et al, 2015) that blend diverse measurements of wind, moisture and temperature and other observations with first guess estimates from model short-term forecasts. While reanalyses effectively reconcile observations with physically-based dynamical models, there are a number of practical problems which result in moisture transport fields typically having substantial systematic time-dependent biases (Trenberth et al, 2011; Robertson et al, 2011; Lorenz and Kunstmann, 2012; Trenberth and Fasullo, 2013; Robertson et al, 2014). The root of the difficulty lies in the fact that model physics (e.g. moist convective parameterizations, turbulence and radiation) each have shortcomings so that assimilating models have biased climatologies. Once initialized, model forecasts (first guesses for analyses) "drift" toward a preferred state that differs from reality. But input data streams that correct this drift are non-stationary in the sense that observing system data densities, and satellite observations especially, have a *time dependent* ability to correct the model first guess fields. Therefore discrete biases develop in water and energy fluxes and transports.

For reanalyses the vertically-integrated atmospheric moisture budget over land grid points is,

$$\frac{\partial \mathrm{W}_a}{\partial t} = \mathrm{VMFC} - \mathrm{P} + \mathrm{ET} + \mathrm{ANA}\,, \qquad (1)$$

i.e., W_a, vapor plus condensate increases as the result of vertically-integrated atmospheric moisture flux convergence (VMFC), evapotranspiration, ET, and is depleted by precipitation, P. In reanalyses, the analysis increment (ANA) represents the departure of the forecast from the analysis

106 divided by the temporal length of the corrector step; specifically in the case of MERRA and
107 MERRA-2, the forcing needed to drive the evolving reanalysis to the final analysis in a 6h
108 corrector step. Ideally this term should be randomly distributed about zero. Instead, time-
109 dependent biases typically characterize ANA, reflecting bias contained in each of the physical
110 terms in (1). Many previous studies (Trenberth and Guillemot, 1998; Lorenz and Kunstmann,
111 2012; Trenberth and Fasullo, 2013) suggest that VMFC estimates have more consistency among
112 reanalyses than P-E derived from the model physics. But still, significant global land trends in
113 VMFC were found by these studies.

114 These issues are seen in Figure 1a which shows reanalysis VMFC monthly anomalies around
115 their respective annual mean. Trends over the period 1979 to 2012 range between -0.03 to 0.08
116 mm d^{-1} $decade^{-1}$ (Table 1) and represent roughly -2.0% to 5.0% of the climatological annual
117 means. These trends are difficult to justify physically given recent estimates of P and ET changes
118 (New et al, 2001; Jung et al, 2010) over recent decades. Discontinuities in the satellite record, in
119 particular with the beginning of Special Sensor Microwave Imager (SSMI) series in July1987, the
120 Advanced Microwave Sounding Unit-A (AMSU A) in late 1998, and Atmospheric Interferometer
121 Sounder (AIRS) in 2002 are known to link with abrupt changes in water and energy fluxes
122 (Bosilovich et al, 2011; Trenberth et al, 2011 and Robertson et al, 2011; 2014; Bosilovich et al,
123 2015).

124 LSMs and other related diagnostic models constrained by observations of precipitation, near-
125 surface atmospheric variables and radiation offer an independent estimate of terrestrial P-ET. In
126 these observationally-constrained models, the water budget has the form

127

128 $$\frac{\partial \mathrm{W}_T}{\partial t} = \mathrm{P} - \mathrm{ET} - \mathrm{RO} \qquad (2)$$

where column terrestrial water (W_T, soil plus vegetation) is sustained by P but depleted by runoff, RO, and ET. Through efforts at the same institutions involved in global data assimilation (e.g. NASA, NCEP and ECMWF) and various internationally coordinated programs (e.g. the Global Land Data Assimilation System, [Rodell et al, 2004]; the Global Soil Wetness Project, GSWP and its successors, GSWP-2 and 3 [Dirmeyer et al, 1999, 2006]; Water and Global Change, WATCH, [Harding et al, 2011]; Trends and Drivers of Regional Sources and Sinks of Carbon Dioxide, TRENDY, [Sitch et al, 2013]), reasonably mature diagnoses of P-ET are now available. These syntheses of land surface state and fluxes have facilitated the quantitative study of droughts and hydrologic variability-- their scale, intensity, and some assessment of changes on a continental and global basis (e.g. Koster et al, 2009, 2011; Wisser et al, 2010; Haddeland et al, 2011; Sheffield et al, 2012; van Dijk et al, 2013). Figure 1b shows corresponding P-ET monthly anomalies from a number of these sources, along with their ensemble mean. To the extent that atmospheric moisture storage anomalies on monthly time scales are small reanalysis VMFC and P-ET should be equivalent. At interannual to near-decadal time scales the agreement between these two quantities is reasonably good with systematic deficits of (excess) moisture transport to land and smaller (larger) P-ET during El Niño (La Niña) events. El Niño events in 1982/1983, 1986/1987, 1991/1992, 1997/1998 coincide with anomalously weak ocean to land moisture transport. After the turn of the century only the 2004/2005 and 2009/2010 events are prominent. At longer scales though the large trends in many of the reanalyses (0.08 and 0.07 mmd^{-1} $decade^{-1}$ for MERRA and CFSR) are not shared by the LSMs whose mean trend is 0.007 mmd^{-1} $decade^{-1}$. Against the prominent interannual and longer excursions the mean LSM global land trend is not significant at the 0.90 level.

152 Further evidence for the likelihood of small trends in these budget components comes from the
153 time series of annual global runoff shown in Figure 2, (Dai, 2009; updated by Dai, 2016). The
154 values shown here are now in terms of mmy^{-1} since the RO values are aggregated over water years
155 (OCT through SEP). The RO trend (0.26 mmy^{-1}) is roughly one half that of P-ET, so that (2)
156 implies a net continental water storage trend over the 30 plus year period. This estimate is likely
157 quite uncertain, although recent work by Reager et al. (2016) finds storage rates of 0.71 (+/-.20)
158 mmy^{-1} over the period 2002-2014 from GRACE measurements. The relevant point here though is
159 that the independent RO and P-ET estimates both provide evidence that large multi-decadal trends
160 in reanalysis VMFC are exaggerated. Taken together with the known inconsistencies introduced by
161 the changing observing system and other observational evidence against such large global trends,
162 the VMFC decadal trends should be treated with considerable skepticism.
163 The objective of this paper is to explore reanalysis VMFC discrepancies with independent
164 LSM-based estimates in more detail. Specifically: (1) We aim to characterize and quantify
165 observing system influences that produce non-physical trends in reanalysis VMFC trend over
166 global land. We explore some of the regional patterns of variability, noting how sensitive global
167 VMFC is to regional uncertainties. (2) In this process, we consider the P-ET record of several
168 observationally-constrained LSMs as a surrogate for validation of VMFC. But uncertainties in
169 forcing data as well as the model formulations are still important error sources (Jimenez et al, 2011
170 and Mueller et al, 2013). Thus, we examine differences among the P-ET estimates and evaluate
171 their utility as a means of reanalysis validation. (3) We then show that using Rotated Empirical
172 Orthogonal Function (REOF) analysis, along with some pre-filtering, artificial steps and trends
173 induced by changing satellite data streams can be largely isolated and removed.
174

2. Data

Our investigation depends primarily on monthly mean data from global reanalyses and observationally constrained land surface models. Since available fields are at different native or archived grid resolutions, we interpolated all data to a 1.0 degree latitude by longitude resolution. Unless otherwise noted, all variability estimates are anomalies that were calculated by removing from the total fields a monthly resolved climatology for the respective data sets at each gridpoint over the period Jan 1979 trough Dec 2010. Some minor departures from these dates are noted in the discussion below.

a. Reanalyses

VMFC is calculated from five state-of-the-art reanalysis projects— the NASA Modern-Era Retrospective analysis for Research and Applications, MERRA, (Rienecker et al. 2011) and an updated version MERRA-2 (Gelaro et al. 2016; Molod et al, 2015; Takacs et al, 2015); the European Centre for Medium Range Weather Forecasting (ECMWF) Interim Reanalysis, ERA-I, (Dee et al. 2011); the new 55-year reanalysis produced by the Japanese Meteorological Agency, JRA-55, (Kobayashi et al, 2015) that extends from 1958 to 2012; and the National Centers for Environmental Prediction Climate Forecast System Reanalysis, NCEP / CFSR (Saha et al. 2010). For ERA-I and JRA55 northward and eastward components of vertically-integrated moisture transport were available and the horizontal flux divergences of these quantities were computed. For MERRA and MERRA-2 the divergence of the vertically-integrated transport was archived as a standard product. CFSR VMFC has been derived by Trenberth et al. (2011) and was obtained directly from the National Center for Atmospheric Research.

196 In addition to the references noted here more in-depth documentation of these reanalysis
197 products, assimilating models and data used can be found at https://reanalyses.org/. Some of the
198 more salient details of these reanalyses are provided in Table 2.

199 *b. Land Surface Models*

200 Products generated by LSMs rely on forcing data from reanalysis output as a first guess but,
201 crucially, incorporate in situ observations, gauge precipitation and some satellite data to relax
202 substantially the biases of these initial estimates (e.g. Dirmeyer et al, 1999; Sheffield et al, 2006;
203 Weedon et al, 2011). Still, there remain uncertainties whose origin and characteristics can be
204 complex; thus, evaluation and validation of ET (and sensible heating and soil moisture) estimates
205 is an ongoing process with assessments that have targeted model formulation and input forcing
206 (Kato et al, 2006; Badgley et al, 2015) and used river discharge, GRACE, and field data (Zaitchik
207 et al, 2010; Rodell et al, 2011) for validation. Intercomparison and validation efforts within larger
208 collaborative efforts of LandFlux (Jimenez et al, 2011; Mueller et al, 2013) and WaterMIP
209 (Haddeland, 2011; Harding et al, 2011) are generating needed quantitative perspectives on flux and
210 surface state uncertainties.

211 Seven different estimates of P-ET are used in this study, each constrained by observational
212 forcing: three are from land surface hydrology models: the Global Land Data Assimilation System
213 initiative, GLDAS-2 (Rodell, 2004); MERRA-Land (Reichle et al, 2011; 2012); the MERRA-2
214 land component (Reichle and Liu, 2014) that runs as part of the assimilation but uses an
215 observation corrected precipitation analysis; and ERA-Interim Land (Balsamo et al, 2015). Two of
216 the models ORCHIDEE (Organising Carbon and Hydrology In Dynamic Ecosystems) model
217 (Kriner et al, 2005) and the Common Land Model, CLM4C, (Oleson et al, 2010; Lawrence et al,
218 2011) are dynamic global vegetation models used in the Trends and Drivers of Regional Sources

219 and Sinks of Carbon Dioxide (TRENDY) initiative (Sitch et al, 2013), a contribution to the
220 REgional Carbon Cycle Assessment and Processes (RECCAP; Canadell et al., 2013). Finally, a
221 diagnostic ET estimate from the MPI-BGC flux data set (Jung et al., 2009; 2010) uses a machine-
222 learning methodology to scale-up eddy covariance measurements from FLUXNET (Baldocchi et
223 al. 2001). A surface energy balance constraint is combined with absorbed photosynthetically
224 active radiation data derived from SeaWiFS (Gobron et al. 2006). For consistency with the
225 derivation of this ET data set we combine it with GPCC V6 precipitation to construct P-ET
226 gridded values. Details and further references for these products are given in Table 3.

227 *c. Other Data*

228 From the CMIP-5 AMIP archive (http://cmip-pcmdi.llnl.gov/cmip5/data_portal.html) we
229 selected five data sets for analysis that roughly sample the breadth of model diversity: the GFDL-
230 HIRAM-C180, GISS-E2-R, HadGEM2-A, MIROC5, and MRI-CGCM. Their monthly P-ET
231 fields are available for the period January 1979 through December 2008. Global runoff as
232 presented by Dai, 2009 (updated by Dai, 2016) were used in Figure 2. These data are used for
233 comparison of time series behavior.

234

235 **3. Isolating non-physical changes in reanalysis VMFC**

236 In the process of generating reanalyses, the analysis increment, ANA, in (1) yields information
237 on the mismatch between the first guess (model forecast) and the observations. Schubert and
238 Chang (1996) employed a least squares analysis of the projection of the physical terms in the
239 moisture budget onto ANA to infer errors in the GEOS reanalysis budget terms. Robertson et al.
240 (2014) used the results of a principal components analysis (PCA) applied to MERRA ANA to
241 regress out artifacts in the water and heat budget flux terms. That study showed that non-physical

modes of variability, due largely to increasing amounts of satellite data and their ability to counter model biases, can have prominent regional to global structure. Since ANA is not readily available for all of the reanalyses used in this study we use PCA in conjunction with pre-filtering to identify these non-physical components in the reanalysis VMFC.

The form of the continental moisture budget equation we analyze is

$$\mathrm{VMFC} = \mathrm{P} - \mathrm{ET} + \mathrm{Res}, \quad (3)$$

where each term is a monthly gridpoint anomaly. Here we have subsumed all uncertainties regarding the moisture storage term and the remaining imbalance into a residual term. This framework differs from (1) since we are taking P-ET from the LSMs, a source independent of the reanalyses. How effectively can we then dissect reanalysis VMFC into its physical part and that due to observing system effects?

Simple PCA provides a compact treatment of variance contributions by mutually orthogonal modes in terms of the spatial coherence of the variability (EOFs) and associated temporal variability (PCs). Successive modes explain the maximum amount of remaining variance. However, individual PCA modes cannot generally be equated with specific sources of variability. Nor can we guarantee physical signals and assimilation artifacts to be collected into separate modes. After examining raw VMFC EOFs and PCs from a PCA decomposition for each reanalysis it was noted that while apparent non-physical variability dominated the leading few modes, these were typically mixed with additional ENSO signals.

To identify the artifacts more clearly we first pre-filtered the VMFC. The first eight PCs of an EOF analysis of Global precipitation Climatology Center (GPCC) precipitation were used to largely remove VMFC physical signals at each grid point via principal component regression:

$$\mathrm{VMFC}^{\mathrm{pf}} = \mathrm{VMFC} - \sum_{i=1}^{m} cov(VMFC, PC_i) \cdot PC_i, \quad (4)$$

where $\mathrm{VMFC}^{\mathrm{pf}}$ is the pre-filtered VMFC and $\mathrm{PC_i}$ is the *ith* PC of the GPCC precipitation. To the extent that ET co-varies with P, we can think of this step as removing VMFC co-varying with P and ET. This approach has two attributes: First, it does not add any source of physical variability to VMFC since by construction it only removes VMFC signals that project onto P variability. This would not be the case if we just subtracted the P anomalies from those of VMFC. Second, since we are first removing much of the physical VMFC signal, this minimizes the likelihood that any subsequent analysis of $\mathrm{VMFC}^{\mathrm{pf}}$ will mistakenly identify physical low-frequency behavior or trends as being artifacts.

We then take a conservative approach of using just the leading few modes of a rotated PCA of $\mathrm{VMFC}^{\mathrm{pf}}$ as representing the bulk of the artifact signals. We chose to rotate the modes (i.e. make linear combinations of them) to collect regional variability into fewer leading modes. Using the Varimax constraint (Richmond, 1986), we rotated the leading 10 modes. The raw PCs were scaled by -1/sqrt(eigenvalues) before input to the rotation matrix so as to preserve orthogonality among the rotated PCs yet relax that constraint for the rotated EOF patterns. The product of the RPC time series and the REOFs recovers each mode's contribution to VMFC variability. The resulting "artifact" modes can then be subtracted from the *raw* VMFC leaving VMFC^{*}, the estimated physical variability,

$$\mathrm{VMFC}^{*} = \mathrm{VMFC} - \sum_{j=1}^{n} REOF_j^{pf} \cdot RPC_j^{pf} \quad , \tag{5}$$

where $\mathrm{REOF_j^{pf}}$ and $\mathrm{RPC_j^{pf}}$ are the *jth* REOF and RPC of $\mathrm{VMFC^{pf}}$. Inserting (5) into (3) we have

$$\mathrm{VMFC} = \mathrm{VMFC}^{*} + \sum_{j=1}^{n} REOF_j^{pf} \cdot RPC_j^{pf} = \mathrm{P} - \mathrm{ET} + \mathrm{Res} \ . \tag{6}$$

Since $\mathrm{VMFC^{pf}}$ is determined by an ad hoc procedure that only minimizes the presence of true physical variability we can't regard this whole signal as being the artifact to be removed. Thus, a more conservative approach is to use only the leading modes that have some obvious relationship to changes in the assimilated data streams. To the extent that VMFC^{*} and P-ET agree, the Res term is explained by the sum of the "artifact" modes and other non-systematic VMFC and P-ET errors. Results of this analysis and the methods to determine the number of modes used along with a sensitivity analysis are presented in section 7.

4. **Regional VMFC and P-ET contributions to global land averages**

a. Regional Interannual Signals and Trends

To determine the extent to which variability shown in Figure 1a, b is manifest regionally, we first examine maps of root mean square (RMS) monthly mean VMFC anomalies (Figure 3, left panel) and trends (right panel) for each reanalysis over the period 1979-2012. Anomalies are departures from the respective monthly varying climatologies. For comparison, the RMS and trend of the ensemble LSM P-ET and reanalysis VMFC are shown in Figure 4. The anomalies were composited before the RMS and trend were calculated. In addition the MERRA and

309 MERRA-2 values were averaged to form one sample so as maintain diversity of the reanalysis
310 systems and LSMs considered.

311 In general terms, the RMS values of ERA-I, JRA-55 and MERRA (Figure 3) agree reasonably
312 with those of the LSM ensemble mean (Figure 4). MERRA-2 and CFSR RMS values are a factor
313 of 2 or more greater in many places over tropical continents. One striking feature is that except for
314 MERRA all reanalyses have larger RMSs over west central Africa compared to that of the LSM
315 ensemble mean. LSM values are typically less than 1.0 mm d^{-1} there whereas reanalysis values
316 exceed 1.5 to 2.0 mm d^{-1} over broad areas and are frequently much larger. There are great
317 observational challenges over tropical continents, especially Africa, not only for the radiosonde
318 density but also for rain gauge and surface atmospheric measurements. Over much of South
319 America, reanalysis VMFC variability agrees well with that of the LSMs, again with MERRA-2
320 and CFSR being much larger. Variability over the headwaters of the Amazon Basin is slightly
321 stronger in the LSM versus the reanalysis ensembles. A separate center of strong variability
322 common to the reanalyses and LSMs (Figure 4a, b) is present over the La Plata Basin, a region of
323 strong convective activity and storm track origination, both modulated by ENSO.

324 Reanalysis VMFC trends show regional structure that does not average out in the ensemble
325 mean and contrasts in many areas with ensemble LSM trends (Figure 3 right, Figure 4c, d). Strong
326 downward trends exceeding 1 $mmd^{-1}decade^{-1}$ dominate Central Africa in all reanalyses except for
327 CFSR which is strongly positive. These signals are much weaker in the LSMs. Upward trends
328 exceeding 1 $mmd^{-1}decade^{-1}$ in the reanalyses are seen in East Africa but are not found in the LSMs.
329 The Maritime Continent and upper reaches of the Amazon basin trend upward in all reanalyses in
330 agreement with the LSMs. Negative VMFC tendencies extend from southern Brazil through the
331 La Plata Basin but are much less organized than the negative P-ET values from the LSMs.

Somewhat surprising is the lack of agreement between the reanalyses in terms of trends over the U.S. In this area of dense observational data LSM trends are near zero but ERA-I and JRA-55 have strong downward trends. In another data rich region over northern and eastern Europe, the upward VMFC trends extend across all reanalyses but are very weak in the LSMs.

One might wonder whether variability of moisture storage Eq. (1) might be large enough to explain some of the discrepancies between LSMs and the reanalyses. We calculated monthly mean $\partial W_a/\partial t$ from MERRA-2 using 1-hour data since this intra-monthly time resolution provides the largest amplitude signal. Monthly mean, atmospheric storage anomalies averaged over global land (not shown) are an order of magnitude less than VMFC and P-ET values shown in Figure 1a, b. However, regional monthly mean RMS W_a tendencies can reach near 0.50 mm/day in subtropical regions and are not negligible for constructing moisture budgets over many areas. Nevertheless, these RMS tendencies are much smaller over tropical continents and cannot explain the reanalysis / LSM discrepancy over Central Africa. Furthermore, regional trends in moisture storage are negligible and cannot explain the differing trends for VMFC compared to P-ET in Figures 3 and 4.

5. **Analysis of regional VMFC errors**

The results of section 4 reinforce our assertion that the differences between the LSM P-ET and the reanalyses VMFC variability on longer than interannual time scale are attributable to systematic reanalysis errors that have largest expression over tropical regions. We now examine several of these specific regions where the LSM P-ET and reanalysis VMFC differ so strongly.

a. Western Equatorial Africa

354 Based on the trend differences in Figures. 3 and 4 we examine the region extending from 10^{o}
355 W to 20^{o}E and from 5^{o}S to 5^{o}N. Time series of VMFC for the reanalyses and mean LSM P-ET is
356 shown in Figure 5a. Each reanalysis shows a distinct change in behavior near the end of 1988.
357 ERA-I drops sharply in a step-like manner as does JRA55. MERRA amplitudes decrease by half
358 but with far less evidence of a change in mean value. Attributing these changes to a specific data
359 stream is difficult. SSMI ingest began in late 1987 for the reanalyses but these effects have to be
360 indirect since the radiances are not used over land. The transition between NOAA9 and NOAA11
361 MSU data also occurs in late 1988. Near the end of 1998 MERRA and JRA55 show a return to a
362 higher amplitude seasonal oscillation with opposite polarity of the pre-1988 period. The MERRA
363 VMFC anomalies also begin a drop over the next 5-7 years. CFSR shows a pronounced increase
364 in 2002 consistent with the beginning of AIRS data. Characteristic of each data set is a change in
365 the annual cycle at the end of 1988 and again in 1998. The annual cycle phase shifts between
366 these periods and its amplitudes decrease in the 1988-1998 period. Clearly, these signals are not
367 physical.
368 Some insight into MERRA's behavior can be gleaned from Figure 5b. In addition to the
369 MERRA VMFC anomalies here we also plot the first two PCs of the global, vertically-integrated
370 moisture analysis increment, ANA, which is the forcing needed to drive the forecast as close as
371 possible to the analysis of observations. These PCs are analogous to those plotted in Figure 7 of
372 Robertson et al. (2011). PC1 minus PC2 also almost exactly recaptures the VMFC time series.
373 The systematic, non-random behavior of the two PCs is evidence of systematic changes in the
374 ability of the data to correct the model forecast first guess. PC1 shows a small but clearly visible
375 drop in late 1987 coincident with SSMI availability. Further distinct drops in late 1998 and again
376 in late 2000 correspond to NOAA15 and NOAA16 AMSU-A data onset. PC2 carries the main

377 signal of the seasonal cycle change. The distinct reduction in seasonal cycle amplitude between
378 1992 and late 1998 corresponds to the tenure of NOAA12. This PC2 behavior is common to all
379 three reanalyses except CFSR suggesting that some combination of satellite-induced changes in
380 the moisture analysis are significant and that the temporal stability of the reanalysis moisture
381 budgets in this region is unreliable.

382 *b. Coastal Ecuador / Colombia*

383 The coastal Ecuador / Colombia region we examined is land area encompassed within the
384 boundaries of Eq. to 10N and 80W to 70W. The choice of this area is motivated by examination of
385 the PCA results to be discussed in section 7 (not shown). Clearly ERA-I is the outlier here (Figure
386 5c), particularly with respect to the jump in VMFC in the spring of 2004. This zero-order change
387 appears associated with the assimilation of METAR surface pressure reports beginning at this time
388 (Figure 11 in Dee et al, 2011). It is unclear as to exactly why these data produce this effect but
389 assimilation of surface pressures that are in disagreement with the first guess forecast could cause
390 analyzed mass changes that subsequently affect the divergent wind, vertical motion and,
391 ultimately, moisture transport fields. These effects can then propagate some distance before being
392 damped. Despite the small area of influence this large near zero-order change in VMFC is the
393 primary signal of ERA-I RPC2. In this region we also see some evidence of jump-like behavior for
394 ERA-I VMFC anomalies in late 1987 although the presence of what appear to be ENSO-related
395 signals in 1982/83 and 1986 complicate the actual magnitude of changes. Why this VMFC
396 increase is more pronounced for ERA-I than in MERRA or JRA55 is not yet clear. Except for this
397 ERA-I problem after NH Spring 2004, the agreement between the LSM mean and the reanalyses is
398 quite good after about 1990.
399

c. Central U.S.

A large downward trend in ensemble mean VMFC over the central U.S. was noted in Figures. 3 and 4c, driven primarily by ERA-I and JRA-55. The time series in Figure 5d shows a distinct downward transition for these two reanalyses of approximately 0.6 mmd^{-1} at the end of 1994. Since VMFC climatological values are of comparable size in this region (Figure 6a) this represents a significant change in the moisture balance. Although there is also a decrease in P-ET from the LSMs before and after 1995, that change is much more gradual. A change in sources of conventional data from historical archives to the ECMWF operational feed beginning in 1995 (Uppala et al, 2005) may explain its distinct VMFC decrease In ERA-I. JRA-55 incorporated this same data as used by ECMWF (Kobayashi et al, 2015) which may explain the similar discontinuity in that system. Complicating this interpretation is the fact that 1995 also marks the transition from NOAA 11 to NOAA 14 sounder coverage. Because of the diurnal drift of the PM satellite equatorial crossing times there was approximately a 3h diurnal cycle difference between these two sensors. Bosilovich et al. (2015) have analyzed VMFC behavior in MERRA but note that this shift occurs near 2000 and may be more related to the assimilation of ATOVS and AIRS data. The exact attribution of the VMFC changes to different data steams remains to be settled; yet, it is clear that even over data rich regions such at the continental U.S. significant VMFC artifacts exist.

6. LSM P-ET signals and uncertainties

Although observationally-constrained LSMs offer a physically consistent estimate of terrestrial water balance, there exist uncertainties stemming not only from model physics formulation but from the quality of the forcing data (Badgley et al, 2015). Precipitation data sets (e.g. GPCC,

Global Precipitation Climatology Project [GPCP] and others) differ in sampling, gauge under-catch, and data quality. More problematic is near-surface meterorology and radiative forcing. These variables are taken from reanalyses too but are bias adjusted using surface observations and satellite radiative fluxes (e.g. Sheffield et al, 2006, Weedon et al, 2011). Given the contrasts noted between reanalysis VMFC and LSM P-ET it is necessary to assess uncertainties in the LSMs to further quantify their credibility vis-à-vis the reanalyses.

Recent work by the LandFlux-EVAL community has highlighted uncertainties in LSMs, diagnostic retrievals and reanalyses. In an initial assessment of flux estimates over the 1993-1995 period, Jimenez et al. (2011) find ET uncertainties of order 0.50 to 0.70 mmd^{-1} relative to an annual mean of about 1.60 mmd^{-1}. Mueller et al. (2013) extend this study in developing a baseline time series of flux estimates including interannual variability and trends. They attribute much of this uncertainty to differences in precipitation forcing used, the influence of water limited ET regimes and interception by vegetation. On the other hand, there also exists a fair degree of sensitivity to model formulation evident when LSMs are run using identical forcing data (Schlosser and Gao, 2009). Lipton et al. (2015) point to ET differences between satellite driven diagnostic approaches and LSMs noting sensitivities to surface parameters and LSM forcing precipitation. Still, Mueller et al. (2013) find realistic interannual variations in ET from composites of these methods are present, including an upward trend between 1989 and 1997 followed by downward ET trend of during the 1998-2005 period. This behavior is most robust for the LSMs and echoes the earlier results of Jung et al. (2010). One common finding from these and other studies is that no single model can be regarded as sufficient and that multiple models with alternative forcing offer the most reliable syntheses of fluxes.

445 LSM ensemble mean climatological mean P-ET patterns (Figure 6a) look very much like
446 precipitation climatologies with large values over the Amazon Basin, Maritime Continent, and S.E.
447 Asia. Storm tracks impinging on the west coasts of N. America and Chile are present. Still,
448 quantitative differences exist among mean P and ET climatological means (not shown). Despite
449 the climatological uncertainties, the global anomalies in Figure 1b show good coherence. To
450 assess more deeply the character of the P-ET anomalies comprising the ensemble estimate we
451 examine two statistical metrics. The mean signal-to-noise ratio (S/N) is defined as S/N
452 $\equiv \sigma^2_{\overline{LSM}} / \overline{\sigma^2_{LSM'}}$ where $\sigma^2_{\overline{LSM}}$ is the square of the ensemble mean monthly P-ET anomaly and
453 $\overline{\sigma^2_{LSM'}}$ is the mean of the individual squared departures of the P-ET anomalies from the ensemble
454 mean monthly anomaly. The S/N diagnostic (Figure 6b) is a local measure of uncertainty among
455 the LSM members in defining P-ET monthly anomalies. Densely populated and gauged areas of
456 the eastern US, Europe and China have systematically high values ranging from 5 to 8. The
457 periphery of Australia and South America also show values in the range of 3 to 5. Deserts (the
458 Sahara, Central Asia, and the interior of Australia) have the lowest values owing to sporadic rain
459 as well as a dearth of gauges. S/N values within key tropical precipitation regimes of Brazil, New
460 Guinea and central Africa are typically 3 or less and likely suffer most directly from insufficient
461 gauge density. These values are for 1.0 degree resolution data and it is important to keep in mind
462 that spatial averaging to coarser resolution of several degrees enhances these numbers
463 significantly.

464 For a more global skill metric we use the anomaly correlation coefficient (ACC) first
465 introduced by Miyakoda et al. (1972):

466 $$ACC \equiv \frac{\langle p_{ij} m_{ij} \rangle}{\left(\langle p_{ij}^2 \rangle \langle m_{ij}^2 \rangle \right)^{1/2}} \qquad (7)$$

467 Here, $p_{i,j}$ are the P-ET anomalies at gridpoint i, j at any given time for any data set and the
468 ensemble mean anomaly is defined by m. The angle brackets denote area-weighted averaging over
469 all gridpoints i, j within the 60^{o} N/S land domain. ACC values measure spatial pattern fidelity as a
470 function of time. Because we lack independent P-ET validation on these scales these diagnostics
471 are more a measure of P-ET sensitivity to input data and model formulation than of accuracy.
472 Results for the seven different LSM P-ET data sets relative to the six-member ensemble mean are
473 given in Figure 6c. (As noted earlier we average the MERRA-Land and MERRA-2 P-ET values as
474 input to the LSM ensemble mean.) MPI-BGC and ORCHIDEE both use GPCC precipitation.
475 ERA-I uses GPCP precipitation which is strongly tied to GPCC gauge data but also differs because
476 of an adjustment to deal with under-catch of gauges. Thus these three precipitation forcing data
477 sets dominate the six member ensemble mean. Values are reasonably stable and generally lie in
478 the range of 0.75 to .95. Experience has shown that values above 0.60 generally are indicative of
479 agreement on the synoptic scale. Thus, from a global coherence perspective, the data sets are
480 similar in their spatial patterns. Sensitivity to the precipitation forcing has a significant influence.
481 Accordingly, MPI-BGC and ORCHIDEE ACC values are each strongly correlated with the
482 ensemble mean; ERA-I Land also shows high correlations. CLM4C ACC values tend to decline in
483 time, especially after 2000. The much smaller number of precipitation gauges used in the CRU
484 TS3.10 product forcing for CLM4C has been shown to lead to a systematic overestimation of
485 precipitation since the mid 1990s (Trenberth et al, 2014; Dai and Zhao, 2016). This appears to
486 influence the lowering CLM4C correlations with time. MERRA-Land and MERRA-2 P-ET
487 values are systematically low although this happens in part because the Climate Prediction Center
488 "Unified" (CPCU) and CPC Merged Analysis of Precipitation (CMAP) precipitation data represent
489 one sample compared to effectively three GPCC forcing sets. There are notable departures though

with MERRA-Land before 1982 and GLDAS-2 Noah in 2006. These periods reflect outliers that originate in the precipitation forcing (CPCU, and the amalgam of data sets that are used for Noah).

ACC calculations performed separately for P and ET (not shown) yielded uniformly high correlations for the former (>0.80) while those for ET averaged between 0.5 and 0.8. The lower ET correlations likely reflect the varied physical formulations among the models and the uncertainties in radiative forcing and near-surface moisture and temperature.

7. Isolating artifacts via REOF analysis (Assessing space / time variability)

To the extent that we believe the mean LSM P-ET trends (Figure 4b) the differences with the VMFC trends in Figures 3 (right hand side) and 4c indicate the regional trend errors or artifacts inherent in the raw reanalyses. In this section we now determine how effectively the RPCA methodology can be used to capture these effects in a few modes.

a. Adjustment effects on regional trends

In section 3 we outlined the methodology of applying RPCA to the quantity $VMFC^{pf}$ in order to identify the leading structures and temporal variability of artificial variability induced by changes in observing system input. For most of the reanalyses a single RPCA mode, n=1, identified the globally-averaged trends characterized as step-like transitions. However, it was found that typically three modes were needed to effectively capture regional trend artifacts. This determination was made by visually inspecting the RPCs and EOFs of each reanalysis $VMFC^{pf}$ and confirming that those modes contained RPC "discontinuities or steps" that coincided with satellite changes such as SSMI, AMSU or AIRS. We also confirmed that these modes made changes that reduced the regional discrepancy between the trend patterns of VMFC in Figure 3 and the ensemble LSM P-ET in Figure 4d. Only the MERRA and MERRA-2 reanalyses showed that

513 additional modes were significant in changing regional trends. Thus, we applied n=5 for MERRA
514 and MERRA-2 and n=3 for the others as constituting the signal of changes induced by evolving
515 assimilation data input. The sensitivity of trend patterns to inherent subjectivity of this selection
516 process is discussed below.

517 This new VMFC estimate for each reanalysis, $VMFC^{*}$, can then be compared to P-ET of the
518 ensemble mean LSMs. Though $VMFC^{*}$ is now not formally independent of P-ET, none of the P
519 variability has modified $VMFC^{*}$. The effects of these potential adjustments are presented in
520 Figure 7. The left panel contains the area-averaged VMFC signal (black) and diagnosed area-
521 average of the artifact that must be removed (red line). The right panel contains the trends in
522 $VMFC^{*}$, the adjusted reanalyses after this artifact signal has been removed. At the bottom of
523 Figure 7 is the ensemble mean trend of $VMFC^{*}$. Note the geometric progression of the color scale.
524 (Recall that we have averaged the MERRA and MERRA-2 data and considered it as one of four
525 reanalysis systems.)

526 It is clear from Figure 7 (left panel, red lines) that $VMFC^{pf}$ collects a large amount of trend or
527 low frequency variability. For example, the impact of AIRS after 2002 in elevating CFSR VMFC
528 is quite apparent. SSMI availability after 1987 changes ERA-I and JRA-55 noticeably. The large
529 ERA-I VMFC jump in 2004 over coastal Ecuador / Colombia (Figure 4c) has a significant global
530 impact. The artifact time series for each reanalysis also has high frequency signals since
531 significant observing system changes such as SSMI, AMSU and AIRS also affect the VMFC
532 annual cycle. After adjusting the reanalyses at each gridpoint the spatial trend patterns (Figure 7,
533 right panel) and the ensemble mean (bottom) are much smaller compared to those in Figure 3 and
534 show significant changes in structure. Over central Africa the amplitudes of $VMFC^{*}$ and LSM P-
535 ET decreases are much more consistent. Both ensemble $VMFC^{*}$ and LSM P-ET (Figure 4d)

trends hint at a tendency for increases in moisture convergence to the south over Zambia / Angola and north over portions of the Sahel. This pattern suggests perhaps increasing annual latitudinal excursion of the ITCZ in these regions. MERRA and CFSR $VMFC^{*}$ no longer have huge upward trends over Australia and the large VMFC increases over East Africa common to all reanalyses have been removed. The Amazon basin shows increased moisture convergence over time with associated reductions over southern Brazil. There remain differences though with the LSMs positioning the P-ET increases over the headwaters region and the reanalyses having the upward $VMFC^{*}$ trends more toward the east. An interesting aspect of this analysis is the relatively small fraction of $VMFC^{pf}$ total variance needed to explained these regional trend artifacts: JRA55 (18.09%), ERA-I (14.15%), MERRA (19.08%), MERRA-2 (15.29%) and CFSR (11.61%). This indicates that the bulk of $VMFC^{pf}$ variability does not project onto regional $VMFC^{pf}$ trends. Each reanalysis' leading mode largely explains the global average land trend and contributes typically about 6%. Since the pre-filtering removes much of the physical signal we interpret this remaining $VMFC^{pf}$ as predominantly error in higher frequency VMFC regional signals.

The global land area-averaged $VMFC^{*}$ trends (60^{o} N/S) shown in Figure 8 are now each much reduced. The ensemble mean value (Figure 8b) is now 0.016 (+/- 0.13) $mmd^{-1}decade^{-1}$ over the period 1979 to 2012. This result is much closer to the mean trend of the LSMs (0.007 [+/-0.010] $mmd^{-1}decade^{-1}$) and the AMIPs (0.012 [+/- 0.016] $mmd^{-1}decade^{-1}$).

Perhaps the most significant outcome of the VMFC adjustment is the improved agreement with LSM P-ET in terms of regional trend patterns and their amplitudes. With these adjustments the pattern correlation of the trends increases from 0.41 to 0.55. Two applications of a 9-point spatial filter were used prior to determining the correlations that raised the correlations by about 0.09. To gauge how sensitive this result is to our selection of modes we repeated the analysis with

559 1-, 2- and 3-mode only corrections. The resulting pattern correlations for the ensemble mean
560 VMFC* with the LSM mean were .048, 0.48 and 0.50, respectively. Higher modes (n > 5) were
561 examined but their relationship to satellite changes was not clear; in keeping with a conservative
562 approach to making corrections these were not used. Individual pattern trend correlations were:
563 JRA-55 = 0.48, ERA-I = 0.61, MERRA = 0.43, MERRA-2 = 0.44 and CFSR = 0.38. Though
564 ERA-I exceeded the ensemble mean, the value for an ensemble with ERA-I removed is 0.52, thus
565 supporting the value of an ensemble strategy as generally providing more skill than individual
566 ensemble members.

567 The agreement in trend patterns and amplitudes between VMFC* and the LSMs (Figure 7,
568 bottom and Figure 4d) is therefore quite improved over the raw VMFC. Studies seeking to explain
569 decadal changes and trend patterns like these have consistently pointed to SST variations as
570 important controls on regional hydrologic anomalies even if details of patterns, seasonality, and
571 intensity remain unresolved. The AMO has been found to influence rainfall over the Sahel
572 (Folland et al, 1986; Giannini et al, 2003), northeast Brazil (Hastenrath and Greisher, 1990;
573 Folland et al, 2001) and the U.S. (Enfield et al, 2001). Gloor et al. (2013) note the effects of
574 Atlantic SST changes on the upward trend in wet season rainfall over Amazonia since 1980. Low
575 frequency ENSO-like behavior of Pacific SSTs (Power et al, 1999; Zhang et al, 1999) has been
576 argued as forcing for Global Monsoon variations (Wang et al, 2012). Positive phases of the PDO,
577 the North Pacific component of this SST variability, are associated with an increase in
578 precipitation in the central and northern parts of the Amazon but decrease in the southern parts
579 (Marengo, 2004). These PDV teleconnections are global as evidenced by Lyon and DeWitt (2012)
580 who have shown that recent Spring declines in East African rainfall are tied to cold eastern tropical
581 Pacific SSTs. Asefi-Najafabady and Saatchi (2016) have noted a continued downward trend in

582 precipitation over central Africa by merging CRU and TRMM data, though Washington et al.
583 (2013) strongly caution against reliance on any precipitation data set in this part of Africa.
584
585 *b. Adjustment effects on interannual variability*
586 Time series of globally-averaged VMFC for the individual corrected reanalyses are given in
587 Figure 8a with Figure 8b showing time series area-averages of ensemble corrected reanalyses and
588 the mean LSMs. Comparing Figure 8a to Figure 1a the reduced trends reveal more consistent
589 interannual VMFC signals among the reanalyses and the relationship between interannual VMFC*
590 and Niño 3.4 SST anomalies is much clearer. Ensemble mean reanalyses and LSMs (Figure 8b)
591 correlate well (cor = 0.86). Ensemble averaging over multiple AMIP experiments reduces internal
592 atmospheric variations (i.e. "weather noise") that cannot represent the correct deterministic signals
593 that were observed (Battisti and Bretherton, 2000). Thus, the remaining AMIP signal is only that
594 component forced by SST. Differences in AMIP P-ET or VMFC anomaly response structure to
595 SST anomalies are also present. These factors lower the AMIP correlation with the LSMs and
596 VMFC* (cor = 0.64 and cor = 0.52, respectively). The agreement between the three data sets thus
597 confirms the significant role that interannual SST variations play in land / ocean moisture
598 exchange.
599 To explore the degree to which spatial VMFC patterns have been affected, we assess the
600 changes in VMFC (equivalently, P-ET) patterns via ACC, this time between the individual
601 corrected reanalyses and the ensemble LSM (Figure 9). The ACC of the ensemble mean raw
602 reanalyses is also plotted (black dotted line) which indicates that on average the adjustment
603 process has slightly degraded VMFC agreement with P-ET (less than 0.05 on average). Here we
604 see that on an individual reanalysis basis, the ACC is typically only 0.35 to 0.60. Again, some of

605 this limitation is local to regional details of the ensemble LSM P-ET values. But the skill of the
606 ensemble mean corrected reanalysis exceeds that of the individual members.
607 An indication of where VMFC and P-ET agreement has changed can be gleaned from local
608 correlations between their time series (Figure 10). There is excellent agreement in locations of
609 dense station sampling and significant rainfall but there is a strong resemblance between Figure
610 10a and the S/N estimates of Figure 6b. This indicates the significant limitations of the LSM
611 signals likely produced by rain gauge sparse density. This lack of station coverage means that in
612 some areas even where the corrected reanalyses are improved the LSMs have such poor ability to
613 discern signals that they cannot confirm this. Changes in correlation with the LSM ensemble P-ET
614 compared to the raw reanalyses is shown in Figure 10b. There are areas of improvement as well as
615 reduced agreement. Many areas in the tropics are improved, but the sparsely gauged areas in
616 Africa and the headwaters of the Amazon are not. In central Asia agreement with the LSMs shows
617 both strong positive and negative changes in VMFC / LSM agreement.
618 The results of Figures 9 and 10 might raise concerns about the effect the REOF adjustments
619 have on interannual signals. To check this we composited VMFC and P-ET anomalies based on
620 warm Nino 3.4 SST anomaly maxima during boreal winters 1982/1983, 1986/1987, 1991/1992,
621 1994/1995, 1997/1998, 2004/2005 and 2009/20010 (not shown). Pattern correlations of VMFC
622 and VMFC* with P-ET were 0.82 and 0.79, respectively. All three composites had regional
623 anomaly patterns the canonical precipitation anomalies first isolated by Ropelewski and Halpert
624 (1987) and more recently by Camberlin et al. (2001), Grimm (2003), Hendon (2003) and Malhi
625 and Wright (2004). We conclude from these results that interannual variability related to SST
626 forcing is not significantly altered by our RPCA adjustments.
627

8. Conclusions

In this study we have sought to characterize the uncertainties in estimating variations in moisture transport of moisture from ocean to land. Reanalyses and LSMs offer two nearly independent methodologies for estimating components of the atmospheric water budget. On seasonal and longer time scales moisture transport should be equivalent to net precipitation minus evaporation. Though reanalyses offer VMFC estimates determined from dynamical modeling constraints on observations, the episodic introduction of new data sources, particularly satellite data streams, has introduced serious time dependent biases. LSMs offer similar physically-based constraints on precipitation and surface meteorological forcing in determining P-ET. Though these forcings also have their own uncertainties our assessment shows that they offer a strong quantitative assessment of VMFC issues. We have also shown that RPCA diagnostics, though ad hoc, can be applied to adjust the raw VMFC estimates. A posteriori, these error reductions are justified by improved agreement of regional trends between the LSM P-ET and VMFC regional trends. Our findings can be summarized as follows:

i. The large trends in near-global mean moisture convergence over land during the period 1979-present in reanalyses are predominantly an artifact due to changes in assimilated data streams and the ability of those data streams to correct model biases. These biases differ among reanalyses due to their differing physical parameterization formulations and aspects of the data assimilation methodology. Averaged over the ensemble LSMs a small net positive trend in P-ET (0.007 $mmd^{-1}decade^{-1}$) is found, but is only significant at 90% confidence.

ii. Corrections to VMFC (P-ET) using RPCA with pre-filtering to identify the non-physical signal are effective in removing many of the problems and substantially enhance the agreement

651 in regional P-ET trends during the 1979-2012 period. RPCA-based adjustments also result in
652 an improvement in trend field correlation from cor=0.41 to 0.55. Simple PCA is likely to fail
653 as signals of the artifacts and those of physical variability are mixed. The decision on how
654 many modes are needed to represent artifact structure in any reanalysis is subjective and
655 depends on cross-referencing RPCA results with assimilated observational data stream
656 metadata.

657 iii. Interannual ENSO-related variations and their decadal-scale modulation are highly
658 consistent between the LSMs and adjusted VMFC time series (cor = 0.86) and composite El
659 Niño P-ET and VMFC patterns (cor = 0.84). Though the adjustments are not needed to detect
660 these interannual signals (Figure 1a), the agreement between VMFC* and P-ET interannual
661 variability is more evident in time series plots (Figure 8b).

662

663 iv. Despite uncertainties inherent with observationally-constrained LSMs, these syntheses can
664 help identify and corroborate more problems associated with reanalysis data changes. The
665 sparseness and uneven sampling of precipitation gauging in remote areas (e.g. tropical
666 continents, especially central Africa) are a significant uncertainty in estimating interannual
667 variability. However, corrections to near-surface meteorology and radiative forcing are
668 important (Ngo-Duc et al, 2005) and need additional scrutiny.

669

670 v. CMIP5 AMIP experiments, despite having somewhat distorted VMFC patterns not directly
671 studied here, also corroborate our estimated VMFC* corrections. Though not encumbered with
672 the effects of changing observing systems, these experiments can only confirm the role of SST

673 forcing since internal atmospheric variability is only at best stochastically consistent with the
674 historical record.
675
676 How do we envision the utility of the present results? The broader problem of reconstructing
677 water and energy budget variability, whether from reanalyses assimilating observations or from
678 diagnostic methods (e.g. Pan et al, 2012; Van Dijk et al, 2014; Rodell et al, 2015; L'Ecuyer et al,
679 2015) requires identifying and accounting for time dependent biases. The present results are a step
680 in that direction in that they would facilitate combining VMFC and P-ET estimates in a diagnostic
681 approach. From a broader perspective additional opportunities are apparent for indirect checks and
682 estimates of VMFC in (1) and (2). Satellite retrievals of W_a, P and ET are one direct means. The
683 accuracy of these retrievals varies according to space and time. Robertson et al. (2014) have
684 recently shown that existing P and especially E estimates over global oceans have serious
685 uncertainties for the purpose of climate variability studies. Though retrieval physics errors
686 contribute to these problems the inter-calibration of multiple sensors and temporal changes in
687 global sampling are also important issues. Improvements are actively being pursued. The lack of
688 robust passive microwave remote sensing before late 1987 on which these retrievals are based is a
689 limitation. From the terrestrial side, eq. (2), it is possible to determine changes in W_T directly via
690 the Gravity Recovery And Climate Experiment (GRACE) satellite mission (Tapley et al., 2004),
691 and with RO measurements from river and streamflow gauges, recover a P-ET estimate. Though
692 GRACE measurements are a unique resource for enabling this approach those data exist only since
693 2003.
694 Another source of information may come from reduced observation reanalyses that in addition
695 to SST, sea-ice and radiative constituents, also assimilate surface pressure observations (Compo et

al, 2011) and marine wind speeds (Poli et al, 2015). While these less robust data constraints also perhaps minimally enforce actual synoptic weather realism, discontinuities in their multi-decadal records appear to be less of a problem than those in the satellite record. These limitations are offset by the property that the SST (and surface pressure) forcing is largely free from the more discrete changes in atmospheric observing systems. Conceivably these integrations could also be run with observed land surface forcing as was applied in MERRA-2.

Ideally, improved model physical parameterizations and removal of data stream biases would mean that analysis increments or innovations would essentially be unbiased and normally distributed. However, model physics improvements (e.g. AGCM convective, turbulence and cloud parameterizations; LSM soil, vegetation and water routing formulations) are long-term development efforts and our discontinuous data streams, particularly from satellites, will always present a time varying capability to correct assimilating model errors in water and energy fluxes. Continued re-evaluation of these modeling, retrieval and in situ resources is necessary to narrow uncertainties in quantifying climate variability.

Acknowledgments

This work was sponsored in large part by the NASA Energy and Water cycle Studies (NEWS) Program, Dr. Jared Entin, Program Manager. The authors wish to acknowledge the institutions and individuals involved in development and production of the many data sets listed here. (For references and data access see tables 1 and 2.) The authors also gratefully acknowledge the many individuals who discussed, enabled data access, and contributed data sets used in this investigation: Drs. Steven Sitch and Sam Levis for access to the CLM4C data under the auspices of the TRENDY initiative, Drs. Graham Weedon and Ben Poulter for the ORCHIDEE integrations from

719 the WFDEI component of the WATCH program. Dr. Markus Reichstein facilitated access to the
720 MPI-BGC satellite-based ET estimates. MERRA and MERRA-2 are distributed by the Goddard
721 Earth Sciences (GES) Data and Information Services Center (DISC). JRA-55 data were obtained
722 from the project website (http://jra.kishou.go.jp/JRA-55/index_en.html). CFSR VMFC fields were
723 obtained from the National Center for Atmospheric Research Climate Analysis Section Data
724 holdings (http://www.cgd.ucar.edu/cas/catalog/newbudgets/index.html). We thank Dr.
725 Aiguo Dai for providing the updated global river runoff time series.
726
727

728 References

729

730 Asefi-Najafabady, S. and Saatchi, S., 2013. Response of African humid tropical forests to recent
731 rainfall anomalies. *Phil. Trans. Roy. Soc. London B: Biol. Sci.*, **368**(1625), 20120306.
732 Balsamo, G., Albergel, C., Beljaars, A., Boussetta, S., Brun, E., Cloke, H., Dee, D., Dutra, E.,
733 Pappenberger, F., de Rosnay, P. and Munoz Sabater, J., 2012. ERA-Interim/Land: A global
734 land-surface reanalysis based on ERA-Interim meteorological forcing, ERA Report Series,
735 ECMWF, Shinfield Park. *Reading*.
736 Balsamo, G., Albergel, C., Beljaars, A., Boussetta, S., Brun, E., Cloke, H., Dee, D., Dutra, E.,
737 Muñoz-Sabater, J., Pappenberger, F., de Rosnay, P., Stockdale, T., and Vitart, F.: ERA-
738 Interim/Land: a global land surface reanalysis data set, *Hydrol. Earth Syst. Sci*., 19, 389-
739 407, doi:10.5194/hess-19-389-2015, 2015.
740 Badgley, G., Fisher, J.B., Jiménez, C., Tu, K.P., Vinukollu, R., 2015. On uncertainty in global
741 terrestrial evapotranspiration estimates from choice of input forcing datasets. *J.*
742 *Hydrometeorol.,* **16**(4): 1449-1455.
743 Bretherton, C. S. and D. S. Battisti, 2000. An interpretation of the results from atmospheric general
744 circulation models forced by the time history of the observed sea surface temperature
745 distribution. *Geophysical Research Letters*, *27*(6), pp.767-770.
746 Bosilovich, M. G. and Schubert, S. D.: Water vapor tracers as diagnostics of the regional
747 hydrologic cycle, *J. Hydrometeorol*., **3**, 149–165,
748 doi:10.1175/15257541(2002)003<0149:WVTADO>2.0.CO;2, 2002.
749 Bosilovich, M. G., F. R. Robertson, and J. Chen, 2011. Global energy and water budgets in
750 MERRA. *J. Climate*, **24**(22),5721-5739.

751 Bosilovich, M. G., J. D. Chern, D. Mocko, F. R. Robertson, and A. M. da Silva, 2015: Evaluating
752 observation influence on regional water budgets in reanalyses. *J. Climate*, **28**, 3631–3649,
753 doi:10.1175/JCLI-D-14-00623.1.
754 Camberlin, P., S. Janicot, and I. Poccard, 2001: "Seasonality and atmospheric dynamics of the
755 teleconnection between African rainfall and tropical sea-surface temperature: Atlantic vs.
756 ENSO." *Int. J. Climatology*,**21**, 8,973-1005.
757 Canadell, J. G., P. Ciais, K. Gurney, C. Le Quéré, S., Piao, M. R., Raupach, and C. L. Sabine,
758 2011: An international effort to quantify regional carbon fluxes. *Eos, Transactions*
759 *American Geophysical Union*, ***92***(10), 81-82.
760 Chou, C., and J. D. Neelin, 2004: Mechanisms of global warming impacts on regional tropical
761 precipitation. *J. Climate*, **17,** 2688–2701.
762 Compo, G. P., and Coauthors, 2011: The Twentieth Century Reanalysis Project. *Quart. J. Roy.*
763 *Meteor. Soc.*, **137**, 1–28, doi:10.1002/qj.776.
764 Dai, A., 2011: Characteristics and trends in various forms of the Palmer Drought Severity Index
765 during 1900–2008. *J. Geophys. Res.: Atmos.*, **116**(D12).
766 _____, 2016:Historical and future changes in streamflow and continental runoff: A review. *AGU*
767 *Monograph. Terrestrial Water Cycle and Climate Change: Natural and Human-Induced*
768 *Impacts*. (Accepted).
769 _____, and T. M. L Wigley, Global patterns of ENSO-induced precipitation. *Geophys. Res.*
770 *Lett.* **27**, 1283–1286 (2000).
771 _____, T. Qian, K. E. Trenberth , J. D. Milliman , 2010: Changes in continental freshwater
772 discharge from 1948 to 2004. *J. Climate*, 22, 2773–2792, doi: 10.1175/2008JCLI2592.1.
773 _____, 2013: The influence of the inter-decadal Pacific oscillation on US precipitation during

1923–2010. *Clim. Dynamics*, **41**(3-4), 633-646.

_____, and T. Zhao, 2016: Uncertainties in historical changes and future projections of drought. Part I: estimates of historical drought changes, *Climatic Change*, (Submitted).

Dee, D.P., with 35 co-authors, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quart. J. R. Meteorol. Soc*., **137**, 553-597 (DOI: 10.1002/qj.828).

Dirmeyer, P. A., A. J. Dolman, and N. Sato, 1999: The Global Soil Wetness Project: A pilot project for global land surface modeling and validation. *Bull. Amer. Meteor. Soc.,* **80,** 851–878.

_____, X. Gao, M. Zhao, Z. Guo, T. Oki and N. Hanasaki, 2006: GSWP-2: Multimodel analysis and implications for our perception of the land surface. *Bull. Amer. Meteor. Soc.*, **87**, 1381–1397. doi: http://dx.doi.org/10.1175/BAMS-87-10-1381.

Ebita, A., S. Kobayashi, Y. Ota, M. Moriya, R. Kumabe, K. Onogi, Y. Harada, S. Yasui, K. Miyaoka, K. Takahashi, H. Kamahori, C. Kobayashi, H. Endo, M. Soma, Y. Oikawa, and T. Ishimizu, 2011: The Japanese 55-year Reanalysis "JRA-55": an interim report, *SOLA*, **7**, 149-152.

Eltahir, E. A. B. and Bras, R. L.: Precipitation recycling in the Amazon Basin, Q. J. Roy. Meteor. Soc., 120, 861–880, doi:10.1002/qj.49712051806, 1994.

Enfield, D. B., Mestas-Nuñez, A. M., & Trimble, P. J. 2001: The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental US. *Geophys. Res. Lett.*, **28**(10), 2077-2080.

Folland, C. K., T. N. Palmer, D. E. and Parker, 1986: Sahel rainfall and worldwide sea temperatures, 1901–85. *Nature*, **320**(6063), 602-607.

797 Folland, C. K., A. W. Colman, D. P., Rowell, and M. K. Davey, 2001: Predictability of northeast
798 Brazil rainfall and real-time forecast skill, 1987-98. *J. Climate*, **14**(9), 1937-1958.
799 Giannini, A., R. Saravanan, and P. Chang, 2003: Oceanic forcing of Sahel rainfall on interannual
800 to interdecadal time scales, *Science*, **302**, 1027–1030.
801 Gelaro, R. and co-authors, 2016: The MERRA-2 Project. In Preparation for J. Climate.
802 Gloor, M.R.J.W., R. J. Brienen, D. Galbraith, T. R. , Feldpausch, J. Schöngart, J. L. Guyot, J. C.
803 Espinoza, J. Lloyd, and O. L. Phillips, 2013: Intensification of the Amazon hydrological
804 cycle over the last two decades. *Geophys. Res. Lett.*, **40**(9),1729-1733.
805 Gobron, N., B. Pinty, F. Mélin, M. Taberner, M. M. Verstraete, M. Robustelli, and J.-L.
806 Widlowski, 2007: Evaluation of the MERIS/ENVISAT FAPAR product. *Adv. Space Res.*,
807 **39**, 105–115, doi:10.1016/j.asr.2006.02.048.
808 Grimm A. M. 2003. The El Nino impact on the summer monsoon in Brazil: regional processes
809 versus remote influences. *J. Climate*, **16**, 263–280.
810 Gu, G., R. F. Adler, G. J. Huffman, and S. Curtis, 2007: Tropical rainfall variability on
811 interannual-to-interdecadal and longer time scales derived from the GPCP monthly
812 product. *J. Climate*, **20**, 4033–4046, doi:10.1175/JCLI4227.1.
813 Greve, P., B. Orlowsky, B. Mueller, J. Sheffield, M. Reichstein, *and* S. I. Seneviratne *(*2014*),*
814 Global assessment of trends in wetting and drying over land*, Nat. Geosci.,* **7***(*10*),* 716–721*,*
815 *doi:*10.1038/ngeo2247.
816 Haddeland, I., et al, 2011: Multimodel estimate of the global terrestrial water balance: Setup and
817 first results. *J. Hydrometeor.*, **12**(5), pp.869-884.

Harding, R., M. Best, E. Blyth, S. Hagemann, P. Kabat, L. M. Tallaksen, T. Warnaars, D. Wiberg, G. P. Weedon, H. van Lanen, F. Ludwig, I. Haddeland, 2011: WATCH: Current knowledge of the terrestrial global water cycle. *J. Hydrometeor.*, **12**, 6, pp. 1149–1156.

Hastenrath, S. and L. Greischar, 1993. Circulation mechanisms related to northeast Brazil rainfall anomalies. *J. Geophys. Res.*, **98**(D3), 5093-5102.

Hendon, H.H., 2003. Indonesian rainfall variability: Impacts of ENSO and local air-sea interaction. *J. Climate*, **16**(11), 1775-1790.

Jiménez, C., and Coauthors, 2011: Global intercomparison of 12 land surface heat flux estimates. *J. Geophys. Res.*, **116**, D02102, doi:10.1029/2010JD014545.

Jung, M., M. Reichstein, and A. Bondeau, 2009: Towards global empirical upscaling of FLUXNET eddy covariance observations: Validation of a model tree ensemble approach using a biosphere model. *Biogeosci. Discuss.*, **6**, 5271–5304, doi:10.5194/bgd-6-5271-2009.

Jung, M., and Coauthors, 2010: Recent decline in the global land evapotranspiration trend due to limited moisture supply. *Nature*, **467**, 951–954, doi:10.1038/nature09396.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang, A. Leetmaa, R. Reynolds, R. Jenne, and D. Joseph, 1996: The NMC/NCAR 40-Year Reanalysis Project". *Bull. Amer. Meteor. Soc.*, **77**, 437-471.

Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, W.-S. Wu and S. Lord, 2010: Introduction of the GSI into the NCEP Global Data Assimilation System. *Wea. and Forecasting*, **24**, 1691–1705, doi: 10.1175/2009WAF2222201.1.

840 Kobayashi, S., Y. Ota, Y. Harada, A. Ebita, M. Moriya, H. Onoda, K. Onogi, H. Kamahori, C.
841 Kobayashi, H. Endo, K. Miyaoka, and K. Takahashi, 2015: The JRA-55 Reanalysis:
842 General specifications and basic characteristics. *J. Meteor. Soc. Japan*, **93**, 5-48,
843 doi:10.2151/jmsj.2015-001.
844 Koster, R. D., M. J. Suarez, A. Ducharne, M. Stieglitz, and P. Kumar, 2000: A catchment-based
845 approach to modeling land surface processes in a general circulation model 1. Model
846 structure. *J. Geophys. Res.*, **105** (D20), 24 809–24 822.
847 _____, Mahanama, T. J. Yamada, G. Balsamo, M. Boisserie, P. Dirmeyer, F. Doblas-Reyes, T.
848 Gordon, Z. Guo, J.-H. Jeong, D. Lawrence, Z. Li, L. Luo, S. Malyshev, W. Merryfield, S. I.
849 Seneviratne, T. Stanelle, B. van den Hurk, F. Vitart, and E. F. Wood, 2009: The
850 contribution of land surface initialization to sub-seasonal forecast skill: First results from
851 the GLACE-2 Project, *Geophys. Res. Lett*., **37**, L02402, doi:10.1029/2009GL041677, 2009.
852 _____, Mahanama, T. J. Yamada, G. Balsamo, A. A. Berg, M. Boisserie, M., Dirmeyer, P. A.,
853 Doblas-Reyes, F. J., Drewitt, G., Gordon, C. T., Guo, J.-H. Jeong, W.-S. Lee, Z. Li, L. Luo,
854 S. Malyshev, W. Merryfield, S. I. Seneviratne, T. Stanelle, B. van den Hurk, F. Vitart, and
855 E. F. Wood, 2011: The second phase of the global land-atmosphere coupling experiment:
856 soil moisture contributions to sub-seasonal forecast skill, *J. Hydrometeorol*., **12**, 805–822,
857 doi:10.1175/2011JHM1365.1.
858 Krinner, G., N. Viovy, N. de Noblet-Ducoudré, J. Ogée, J. Polcher, P. Friedlingstein, P. Ciais, S.
859 Sitch, and I. C. Prentice, 2005: A dynamic global vegetation model for studies of the
860 coupled atmosphere-biosphere system, *Global Biogeochem. Cy*., **19**, GB1015,
861 doi:10.1029/2003GB002199.
862 Lawrence, D. M., K. W. Oleson, K., M. G. Flanner, P. E. Thornton, S. C. Swenson, P. J.

863 Lawrence, X. Zeng, Z.-L. Yang, S. Levis, K. Sakaguchi, G. B. Bonan, and A. G. Slater,
864 2011: Parameterization improvements and functional and structural advances in version 4
865 of the Community Land Model, *J. Adv. Model. Earth Syst*., **3**,
866 M03001,doi:10.1029/2011MS000045.
867 L'Ecuyer, T. S., H. K. Beaudoing, M. Rodell, W. Olson, B. Lin, S. Kato, C. A. Clayson, E.Wood,
868 J. Sheffield, R. Adler, G. Huffman, M. Bosilovich, G. Gu, F. Robertson, P. R.Houser, D.
869 Chambers, J. S. Famiglietti, E. Fetzer, W. T. Liu, X. Gao, C. A. Schlosser, E.Clark, D. P.
870 Lettenmaier, K. Hilburn, 2015: The Observed State of the Energy Budgetin the Early 21
871 Century, 2014. *J. Climate*, **28**(21), 8319-8346. doi: http://dx.doi.org/10.1175/JCLI-D-14-
872 00556.1.
873 Lipton, A.E., P. Liang, C. Jiménez, J. L. Moncet, F. Aires, C. Prigent, R. Lynch, J. F.
874 Galantowicz, R. P. d'Entremont, and G. Uymin, 2015: Sources of discrepancies between
875 satellite-derived and land surface model estimates of latent heat fluxes. *J. Geophys. Res.:*
876 *Atmos.*, **120**(6), pp.2325-2341.
877 Lorenz, C., and H. Kunstmann, 2012: The hydrological cycle in three state-of-the-art reanalyses:
878 intercomparison and performance analysis. *J. Hydrometeor*, **13**, 1397–1420.doi:
879 10.1175/JHM-D-11-088.1.
880 Lyon, B., and D. G. DeWitt, 2012: A recent and abrupt decline in the East African long rains.
881 *Geophys. Res. Lett*., **39**, L02702,doi:10.1029/2011GL050337.
882 Malhi, Y. and J. Wright, 2004: Spatial patterns and recent trends in the climate of tropical
883 rainforest regions. *Phil. Trans. Roy. Soc. London B: Biol. Sci.*, **359**(1443), 311-329.
884 Marengo, J. A., 2004. Interdecadal variability and trends of rainfall across the Amazon basin.
885 *Theoretical and applied climatology*, **78**(1-3), 79-96.

McGregor, S., A. Timmermann, M. F. Stuecker, M. H. England, M. Merrifield, F.-F. Jin, and Y. Chikamoto, 2014: Recent Walker circulation strengthening and Pacific cooling amplified by Atlantic warming, *Nat. Clim. Change*, **4**, 888–892, doi:10.1038/NCLIMATE2330.

Molod, A., L. Takacs, M. Suarez, and J. Bacmeister, 2015: Development of the GEOS-5 atmospheric general circulation model: evolution from MERRA to MERRA2. *Geoscientific Model Development*, **8**(5), 1339-1356.

Mueller, B., and Coauthors, 2013: Benchmark products for land evapotranspiration: LandFlux-EVAL multi-dataset synthesis. *Hydrol. Earth Syst. Sci. Discuss.*, **10**, 769–805, doi:10.5194/hessd-10-769-2013.

Miyakoda, K., G. D. Hembree, R. F. Strickler, I. Shulman, 1972: Cumulative results of extended forecast experiments. 1. Model performance for winter cases. *Mon. Wea. Rev.*,**100**, 836-855.

New, M. M. Todd, M. Hulme, and P. Jones, 2001: Precipitation measurements and trends in the twentieth century. *Int. J. Climatol.*, **21**: 1889–1922. doi:10.1002/joc.680.

Ngo-Duc, T., J. Polcher, *and* K. Laval, 2005: A 53-year forcing data set for land surface models, *J. Geophys. Res.,* **110**, D06116, doi:10.1029/2004JD005434.

Oleson, K. W., D. M. Lawrence, G. B. Bonan, M. G. Flanner, E. Kluzek, P. J. Lawrence, S. Levis, S. C., Swenson, Thornton, P. E., Dai, A., Decker, M., Dickinson, R., Feddema, J., Heald, C. L., F. Hoffman, J.-F. Lamarque, N. Mahowald, G.-Y. Niu, T. Qian, J. Randerson, S. Running, K. Sakaguchi, A. Slater, R. Stockli, A. Wang, Z.-L. Yang, and , 2010: Technical description of version 4.0 of the Community Land Model (CLM), NCAR Technical Note NCAR/TN-478+STR, 257 pp., 2010.

908 Onogi, K., J. Tsutsui, H. Koide, M. Sakamoto, S. Kobayashi, H. Hatsushika, T. Matsumoto, N.
909 Yamazaki, H. Kamahori, K. Takahashi, S. Kadokura, K. Wada, K. Kato, R. Oyama, T.
910 Ose, N. Mannoji and R. Taira, 2007: The JRA-25 Reanalysis. *J. Meteorol. Soc. Japan*, **85**:
911 369–432.
912 Pan, M., A. K. Sahoo, T. J. Troy, R. K. Vinukollu, R.K., J. Sheffield, and E. F. Wood, 2012:
913 Multisource estimation of long-term terrestrial water budget for major global river basins.
914 *J. Climate*, **25**(9), pp.3191-3206.
915 Poli, P., H. Hersbach, P. Berrisford, D.P. Dee, A. Simmons and P. Laloyaux, 2015: ERA-20C
916 Deterministic. *ERA Report Series 20*.
917 Power, S., T. Casey, C. Folland, A. Colman, and V. Mehta, 1999: Inter-decadal modulation of the
918 impact of ENSO on Australia. *Clim. Dynamics*. **15**,319_324.
919 Reager, J. T., A. S. Gardner, J. S. Famiglietti, D. N. Weiss, A. Eicker and M. H. Lo: 2016: A
920 decade of sea level rise slowed by climate-driven hydrology. *Science,* **12**;351(6274):699-
921 703. doi: 10.1126/science.aad8386.
922 Reichle, R. H., R. D. Koster, G. J. M. De Lannoy, B. A. Forman, Q. Liu, S. P. P. Mahanama, and
923 A. Toure, 2011: Assessment and enhancement of MERRA land surface hydrology
924 estimates. *J. Climate*, **24**, 6322-6338, doi:10.1175/JCLI-D-10-05033.1.
925 _____, R. H., and Q. Liu, 2014. Observation-Corrected Precipitation Estimates in GEOS-5.
926 *NASA/TM–2014-104606*, Vol. **35.**
927 Rienecker, M., M. J. Suarez, R. Gelaro, R. Todling, J. Bacmeister, E. Liu, M. G. Bosilovich, S. D.
928 Schubert, L. Takacs, G.-K. Kim, S. Bloom, J. Chen, D. Collins, A. Conaty, A. da Silva, W.
929 Gu, J. Joiner, R. D. Koster, R. Lucchesi, A. Molod, T. Owens, S. Pawson, P. Pegion, C. R.
930 Redder, R. Reichle, F. R. Robertson, A. G. Ruddick, M. Sienkiewicz, and J. Woollen,

2011: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications**.** *J. Climate*, **24**, 3624-3648, doi:10.1175/JCLI-D-11-00015.1.

Robertson, F. R., M. G. Bosilovich, J. Chen, and T. L. Miller, 2011: The Effect of Satellite Observing System Changes on MERRA Water and Energy Fluxes. *J. Climate*, **24**, 5197–5217. doi: 10.1175/2011JCLI4227.1

_____, M. G. Bosilovich, J. B. Roberts, R. H. Reichle, R. Adler, L. Ricciardulli, W. Berg, and G. J. Huffman, 2014: Consistency of Estimated Global Water Cycle Variations over the Satellite Era. *J. Climate*, **27**, 6135–6154.

Rodell, P. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C. Meng, K. Arsenault, B. Cos- grove, J. Radakovich and M. Bosilovich, 2004: The global land data assimilation system, *Bull. Amer. Meteorol. Soc.*, **85**, 381–394.

_____, E. B. McWilliams, J. S. Famiglietti, H. K. Beaudoing, and J. Nigro, Estimating evapotranspiration using an observation based terrestrial water budget, *Hydrol. Proc*., **25**, 4082-4092, 2011.

_____, H.K. Beaudoing, T.S. L'Ecuyer, W.S. Olson, J.S. Famiglietti, P.R. Houser, R.Adler, M.G. Bosilovich, C.A. Clayson, D. Chambers, E. Clark, E.J. Fetzer, X. Gao, G. Gu, K. Hilburn, G.J. Huffman, D.P. Lettenmaier, W.T. Liu, F.R. Robertson, C.A. Schlosser, J. Sheffield, E.F. Wood, 2015: The observed state of the water cycle in the early 21st century. *J. Climate*, **28**(21), pp.8319-8346.

Ropelewski, C. F. and M. S. Halpert, 1987: Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation. *Mon. Wea. Rev.,* **115**(8), 1606-1626.

Saha, S. and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015–1057. doi: http://dx.doi.org/10.1175/20.

954 Schubert, S., and Y. Chang, 1996: An objective method for inferring sources of model error. *Mon.*
955 *Wea. Rev.*, **124**, 325–340. 10BAMS3001.1.
956 Schlosser, C., and X. Gao, 2009: Assessing evapotranspiration estimates from the Global Soil
957 Wetness Project Phase 2(GSWP-2), Tech. Rep. 179, MIT Joint. Program. on the Sci. and
958 Policy of Global Change, Cambridge, Mass.
959 Sheffield, J., G. Goteti, and E. F. Wood, 2006: Development of a 50-yr high-resolution global
960 dataset of meteorological forcings for land surface modeling, *J. Climate*, **19** (13), 3088-
961 3111.
962 Sheffield, J., E. F. Wood, and M. L. Roderick, 2012: Little change in global drought over the past
963 60 years. *Nature*, **491**(7424), 435-438.
964 Sitch S, P. Friedlingstein, N. Gruber, S. Jones, G. Murray-Tortarolo, A. Ahlstrom, S. C. Doney, H.
965 Graven, C. Heinze, C. Huntingford, S. Levis, P. E. Levy, M. Lomas, B. Poulter, N. Viovy,
966 S. Zaehle, N. Zeng, A. Arneth, G. Bonan, L. Bopp, J. G. Canadell, F. Chevallier, P. Ciais,
967 R. Ellis, M. Gloor, P. Peylin, S. Piao, C. Le Quere, B. Smith, Z. Zhu, and R. Myneni,
968 2013: Trends and drivers of the regional-scale sources and sinks of carbon dioxide over the
969 past two decades. *Biogeosciences Discussions*, **10**:20113-20177.
970 Sutton, R. T. and D. L. Hodson, 2005: Atlantic Ocean forcing of North American and European
971 summer climate. *Science*, **309**(5731),115-118.
972 Takacs, L. L., M. J. Suarez, and R. Todling, 2015: Maintaining atmospheric mass and water
973 balance in reanalyses. *Quart. J. Roy.* Meteorol. Soc. 00: 1–9, DOI: 10.1002/qj.2786.
974 Tapley, B. D., S. Bettadpur, J. C. Ries, P. F. Thompson, and M. M. Watkins, 2004: GRACE
975 measurements of mass variability in the Earth system. *Science*, **305**(5683), pp.503-505.
976 Taylor, K. E., R. J. Stouffer, G. A. Meehl, 2012: An Overview of CMIP5 and the experiment

977 design." *Bull. Amer. Meteor. Soc.*, **93**, 485-498, doi:10.1175/BAMS-D-11-00094.1.
978 Ting, M., Y. Kushnir, R. Seager, *and* C. Li, 2011: Robust features of Atlantic multi-decadal
979 variability and its climate impacts, *Geophys. Res. Lett.,* **38**, L17705,
980 *doi:*10.1029/2011GL048712.
981 Trenberth, K. E., 1999: Atmospheric moisture recycling: role of advection and local evaporation,
982 *J. Climate*, **12**, 1368–1381,doi:10.1175/1520 0442(1999)012<1368:AMRROA>2.0.CO;2.
983 _____, and C. J. Guillemot, 1998: Evaluation of the atmospheric moisture and hydrological
984 cycle in the NCEP/NCAR reanalyses." *Climate Dynamics* **14**.3,213-231.
985 _____, J. T. Fasullo, and J. Mackaro, 2011: Atmospheric moisture transports from ocean to
986 land and global energy flows in reanalyses. *J. Climate*, **24**(18), 4907-4924.
987 _____, and J. T. Fasullo, 2013: Regional energy and water cycles: Transports from ocean to
988 land. *J. Climate*, **26**(20), 7837-7851.
989 _____, A. Dai, G. van der Schrier, P. D. Jones, J. Barichivich, K. R. Briffa, and J. Sheffield,
990 2014: Global warming and changes in drought. *Nature Climate Change*, **4**, 17-22,
991 doi:10.1038/NCLIMATE2067.
992 Uppala, S. M., and Coauthors, 2005: The ERA-40 Re-Analysis. *Quart. J. Roy. Meteor. Soc.*, **131**,
993 2961–3012, doi:10.1256/qj.04.176.
994 Van Dijk, A. I. J. M., L. J. Renzullo, Y. Wada, and P. Tregoning, 2014: A global water cycle
995 reanalysis (2003–2012) merging satellite gravimetry and altimetry observations with a
996 hydrological multi-model ensemble. *Hydrol Earth Syst Sci* **18**:2955–2973. doi:10.5194/
997 hess-18-2955-2014.
998 Wang, B., J. Liu, H. J. Kim, P. J. Webster, and S. Y. Yim, 2012: Recent change of the global
999 monsoon precipitation (1979–2008). *Clim. Dynamics*, **39**(5),1123-1135.

1000 Washington, R., R. James, H. Pearce, W. M. Pokam and W. Moufouma-Okia, 2013: Congo Basin
1001 rainfall climatology: can we believe the climate models? *Phil. Trans. Roy. Soc., B*,
1002 **368**(1625 (20120296)): 1-7
1003 Weedon, G. P., S. Gomes, P. Viterbo, W. J. Shuttleworth, E. Blyth, H. Österle, J. C. Adam, N.
1004 Bellouin, O. Boucher M. and Best, 2011: Creation of the WATCH Forcing data and its use
1005 to assess global and regional reference crop evaporation over land during the twentieth
1006 century. *J. Hydrometerol.* **12**, 823-848, doi: 10.1175/2011JHM1369.1
1007 _____, S. Gomes, G. Balsamo, M. J. Best, N. Bellouin and P. Viterbo, 2012: README file for
1008 the "WFDEI" dataset (version: 18th September 2013)**.** *(This document is available for*
1009 *download at: www.eu-watch.org/data_availability)*
1010 Wisser, D., B. M. Fekete, C. J. Vörösmarty, and A. H. Schumann, 2010: Reconstructing 20th
1011 century global hydrography: A contribution to the Global Terrestrial Network-Hydrology
1012 (GTN-H). *Hydrol. Earth Syst. Sci.*, **14**, 1–24, doi:10.5194/hess-14-1-2010
1013 Zaitchik, B. F., M. Rodell, and F. Olivera, 2009: Evaluation of the Global Land Data Assimilation
1014 System using global river discharge data and a source to sink routing scheme, 2010: *Water.*
1015 *Resour. Res.*, **46**, W06507, doi:10.1029/2009WR007811.
1016 Zhang, Y., J. M. Wallace, and D. S. Battisti, 1997: ENSO-like inter-decadal variability: 1900–93.
1017 *J. Climate*, **10**, 1004–1020.

Table 1. Trend statistics ($mmd^{-1}decade^{-1}$) for VMFC* over land for various reanalyses and P-ET for LSM members over the period 1979 to 2012. In parentheses are errors calculated using lag-one statistics to account for serial autocorrelation.

Land Surface Model	**MERRA-Land**	**MERRA-2**	**ERA-I Land**	**GLDAS NOAH**	**ORCHIDEE**	**CLM4C**	**MPI-BGC**
P-ET Trend	0.023 (+/-0.019)	-0.002 (+/- 0.019)	-0.001 (+/- 0.017)	0.022 (+/-0.012)	0.009 (+/- 0.017)	0.023 (+/0.015)	0.028 (+/-0.022)
Reanalysis VMFC* Trend	**MERRA**	**MERRA-2**	**ERA-I**	**JRA-55**	**CFSR**		
	0.073 (+/- 0.026)	0.003 (+/- 0.026)	0.081 (+/- 0.020)	-0.030 (+/- 0.018)	0.074 (+/- 0.025)		

1024 Table 2. Summary characteristics of reanalysis data sets used in this study.
1025

	Attributes	References	Comments
ERA-I	Cy31r2, 2006 IFS 80 km (T255 spectral) grid with 60 vertical levels. Jan1979-Dec2012.	Dee et al. 2011; Dee and Uppala 2008; 2009; Simmons et al, 2010	4-DVar system with adaptive estimation of satellite bias correction. RTTOV radiation operator, many revised analysis and physics improvements over ERA-40 (e.g. humidity, O3). http://apps.ecmwf.int/datasets/
JRA55	JMA Dec2009 Operational System. TL319L60 (~55km). Jan1958- Dec2012.	Kobayashi et al, 2015; Ebita et al, 2011	Many updates over JRA25 including a new radiation scheme, 4D-Var,Variational Bias Correction for satellite radiances, varying greenhouse gases. http://jra.kishou.go.jp/comm/application_en.html
MERRA	GEOS-5.2.0 AGCM 0.5° x 0.667° grid; 72 vertical levels. Jan1979-Dec2010.	Rienecker et al.2011; Trenberth et al, 2011; Bosilovich et al, 2011; Robertson et al, 2011.	3D-Var Gridpoint Statistical Interpolation Scheme (GSI) with incremental Analysis Update (IAU). http://disc.sci.gsfc.nasa.gov/uui/search/%22MERRA-2%22
MERRA-2	GEOS-5.12.4 AGCM 0.5° x 0.667° grid ; 72 vertical levels. Jan1980-Dec2012.	Molod et al, 2015; Takacs et al, 2015; Kleist et al. 2009b	GSI, IAU; Significant additional new satellite data assimilated; Now conserves dry air mass. http://disc.sci.gsfc.nasa.gov/daac-bin/FTPSubset.pl?LOOKUPID_List=MAIMCPASM
CFSR	GFS 2009 T382L64 coupled atmosphere-ocean-land surface-sea ice system, Jan1979-Dec2009.	Saha et al. 2010; Wang et al, 2010; Trenberth and Fasullo, 2013	3DVar GSI system, AER Radiation, Noah LSM with MOM ocean model Ocean is 0.25° at the equator, extending to a global 0.5° beyond the tropics, with 40 levels. http://www.cgd.ucar.edu/cas/catalog/reanalysis/index.html

1026
1027
1028

Table 3. Observationally constrained Land Surface Models providing P and ET in this study.

Model	Attributes	Forcing	Comments and references
CLM4C	Five primary sub-grid land cover types (glacier, lake, wetland, urban, vegetated) in each grid cell. The vegetated portion of a grid cell is divided into patches of plant functional types with separate energy and water calculations.	Based on a merged product of Climate Research Unit (CRU) observed monthly 0.5 analysis (v3.0, 1901–2009; New et al, 2000) and the high temporal fidelity NCEP reanalysis forcing.	Land model for the National Center for Atmospheric Research (NCAR) Community Earth System Model and the Community Atmosphere Model. (Oleson et al, 2010; Lawrence et al, 2011; Weedon et al, 2011). http://www.cesm.ucar.edu/models/clm/
ERA-I Land	Hydrology-Tiled ECMWF Scheme for Surface Exchanges over Land (HTESSEL). 80 km res with 3h integration steps.	ERA-Interim near surface meteorology and radiation. ERA-I precipitation is rescaled using GPCP v2.1.	Updated physics including soil hydrology, a new snow scheme, multi-year satellite based vegetation climatology. (Balsamo et al, 2012, 2015). http://www.ecmwf.int/en/research/climate-reanalysis/era-interim/land
GLDAS-2 Noah	1-D column model which can be executed in either coupled or uncoupled mode. Governing equations of the physical processes of the soil-vegetation-snowpack medium.	Updated Sheffield et al. (2006) forcing dataset based on the NCEP–NCAR reanalysis near-surface meteorological variables. GPCP, TRMM precip, SRB Radiation. CRU meteorological data used to correct biases.	Development began1993 through a collaboration of investigators from public and private institutions, spearheaded by the National Centers for Environmental Prediction. (Chen et al, 1996; Koren et al. 1999; Rodell, 2004.) http://ldas.gsfc.nasa.gov/gldas/
MERRA-Land	GEOS-5 Six-layer catchment land surface model Koster et al. 2000. The basic land surface element, or "tile," is a topographically determined hydrological catchment non-congruent with overlying MERRA lat x lon grid	Precipitation forcing with the global gauge-based NOAA Climate Prediction Center "Unified" (CPCU) product.	Generated by offline version of the land component of the MERRA system. Revised parameter values in the rainfall interception model from those in the MERRA surface meteorological forcings. (Reichle et al, 2011; 2012). http://gmao.gsfc.nasa.gov/research/merra/merra-land.php
MERRA-2	Updated version of catchment model used in MERRA.	Precipitation constraints comprised of anomalies from CMAP V0011 and RT pentad product plus GPCP v2.1 climatology. Near surface meteorology and radiation from MERRA-2.	MERRA 0.5 deg, hourly time series of precipitation (background) are constrained to have the same daily totals as constraining CMAP/GPCP data. (Reichle and Liu. 2014). http://gmao.gsfc.nasa.gov/reanalysis/MERRA-2
MPI-BGC	Machine-learning methodology, "model tree ensembles", to up-scale eddy covariance (EC) measurements from FLUXNET (Baldocchi et al. 2001) to a 0.5 degree monthly product	AVHRR NDVI data SeaWiFS fPAR; CRU near-sfc temperature	ET estimates use GPCC V6 precip in classification step (Jung et al., 2009; 2010.). We thus use GPCC precip to make consistent P- ET. https://www.bgc-jena.mpg.de/bgi/index.php/
ORCHIDEE	Solves water-energy-carbon budget. Represents ecosystem in terms of a range of Plant Functional Types using big leaf approach. Computes its own phenology.	ERA-I fields constrained with CRU 3.2 Sfc air temp and GPCC v6 monthly precip Forcing precipitation, air temperature, wind, solar radiation, humidity and atmospheric CO_2	SECHIBA land-surface scheme, which is dedicated to the surface energy and water balances, and the carbon and vegetation model STOMATE. Kriner et al, 2005; Weedon et al. (2012); Sitch et al, 2013; Poulter (personal communication). http://forge.ipsl.jussieu.fr/orchidee/wiki

1033

1034 Figure Captions

1035

1036 Figure 1. (*a*) Time series of global land area-average (60^{o}N/S) vertically-integrated moisture
1037 convergence anomalies (VMFC) from various reanalyses. Units are mmd^{-1} (left hand side scale).
1038 Niño 3.4 SST anomalies (^{o}C) are shaded gray with inverted scale on the right hand side. (*b*) Same
1039 as above but for P-ET from individual LSMs. Ensemble mean is shown in black. Units are mmd^{-}
1040 ^{1}. For both time series a three month running smoother has been applied for display purposes.

1041

1042 Figure 2. Annual mean anomalies of LSM ensemble mean P-ET and globally-integrated
1043 streamflow from Dai (2016). Units are mm yr^{-1}. LSM monthly anomalies from Figure 1 have been
1044 summed over water year intervals beginning in October 1979.

1045

1046 Figure 3. Statistics for monthly mean VMFC anomalies for various reanalyses over the period
1047 1979-2012. *Left (a-e):* RMS of deviations (mmd^{-1}). *Right (f-j):* Trends (mmd^{-1}decade^{-1}).
1048 Figure 4. Ensemble mean statistics for reanalysis and LSM monthly anomalies over the period
1049 1979-2012. *Left panel*: RMS (mmd^{-1}) . *Right Panel:* Trends (mmd^{-1}decade^{-1}).

1050

1051 Figure 5. Time series of reanalyses VMFC and ensemble LSM anomalies (mm d^{-1}) over (a)
1052 Equatorial Africa (c) Coastal Colombia / Ecuador region, and (d) Central U.S. (c) Shows MERRA
1053 VMFC (red) and first two global PCs of vertically-integrated moisture increment.

1054

1055 Figure 6. (*a*) Ensemble mean climatological P-ET (mmd^{-1}). (*b*) S / N for LSMs (see text for details
1056 on calculation). (*c*) P-ET ACC time series of each LSM with the ensemble mean.

1057

1058 Figure 7. (*Left*) Time series of globally averaged, 60^{o} N/S, reanalysis VMFC (black) and area-
1059 averaged corrections (red). Niño 3.4 SST x 0.1 is plotted as gray shading with inverted scale on
1060 right hand side. (*Right*) Trends (mmd^{-1} decade^{-1}) in VMFC* over the period 1979-2012 for
1061 various reanalyses after corrections have been applied. Compare to Figure 3, *right* which are the
1062 uncorrected trends. (Bottom) Ensemble mean corrected VMFC* trends.

1063

1064 Figure 8. (*a*) Time series of individual corrected reanalysis VMFC* global land area-average
1065 (60^{o}N/S). (*b*) P-ET from individual LSMs (black) and mean VMFC* from corrected reanalyses
1066 (red) and AMIPs (cyan). A three-month running smoothing is applied. Units are mmd $^{-1}$. Niño
1067 3.4 SST anomalies are plotted in gray shading with inverted scale (deg C) on right.

1068

1069 Figure 9. Anomaly correlations, ACC, between individual corrected reanalysis VMFC* and
1070 ensemble mean LSM P-ET. ACC for the ensemble corrected and uncorrected reanalyses with the
1071 ensemble LSMs is shown by the solid (dotted) black lines, respectively. A running 3-month
1072 smoother has been applied to each time series for display.

1073

1074 Figure 10. Local correlations (1979-2012): (a) Ensemble mean adjusted reanalysis VMFC* and
1075 ensemble mean LSM P-ET, (b) Adjusted VMFC* correlation with LSMs minus raw reanalysis
1076 ensemble reanalyses VMFC correlation with ensemble LSM P-ET. Note different color scales.

1077

1078

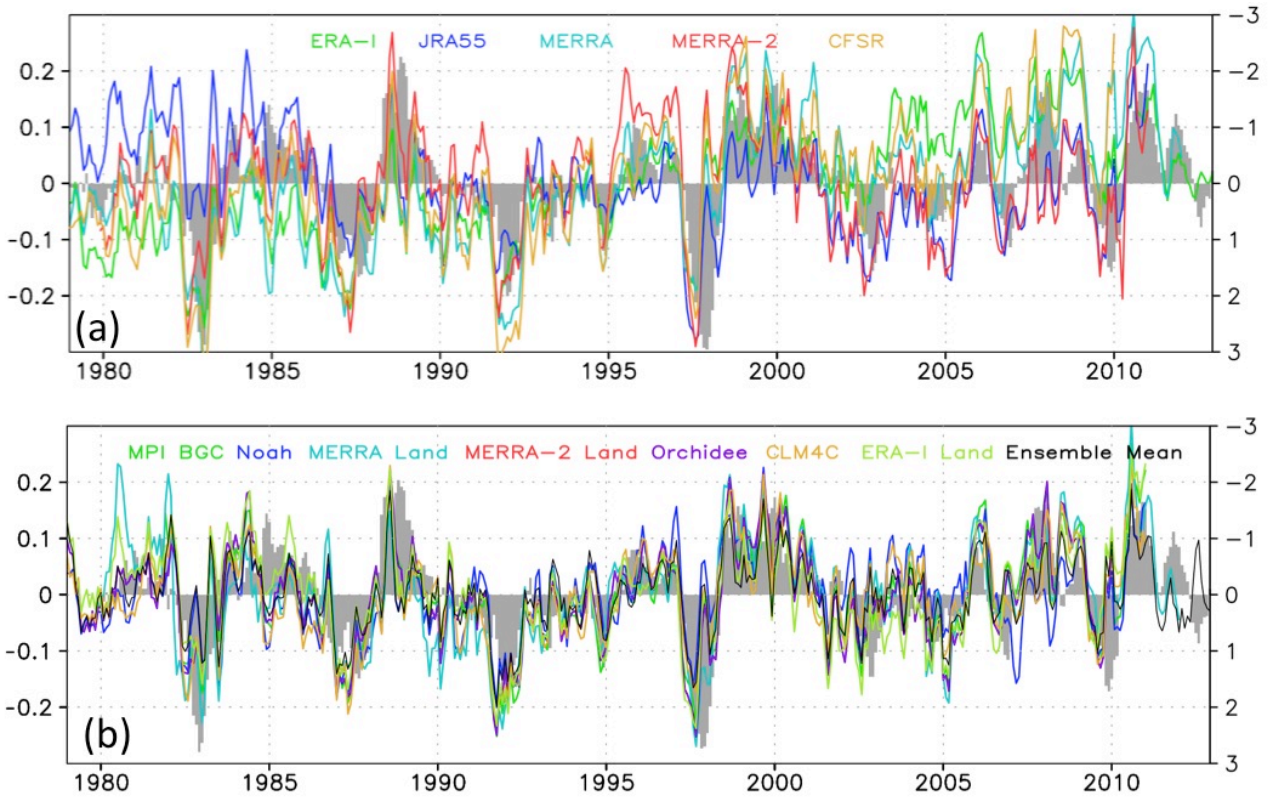


Figure 1. (*a*) Time series of global land area-average (60°N/S) vertically-integrated moisture convergence anomalies (VMFC) from various reanalyses. Units are mmd^{-1} (left hand side scale). Niño 3.4 SST anomalies (°C) are shaded gray with inverted scale on the right hand side. (*b*) Same as above but for P-ET from individual LSMs. Ensemble mean is shown in black. Units are mmd^{-1}. For both time series a three month running smoother has been applied for display purposes.

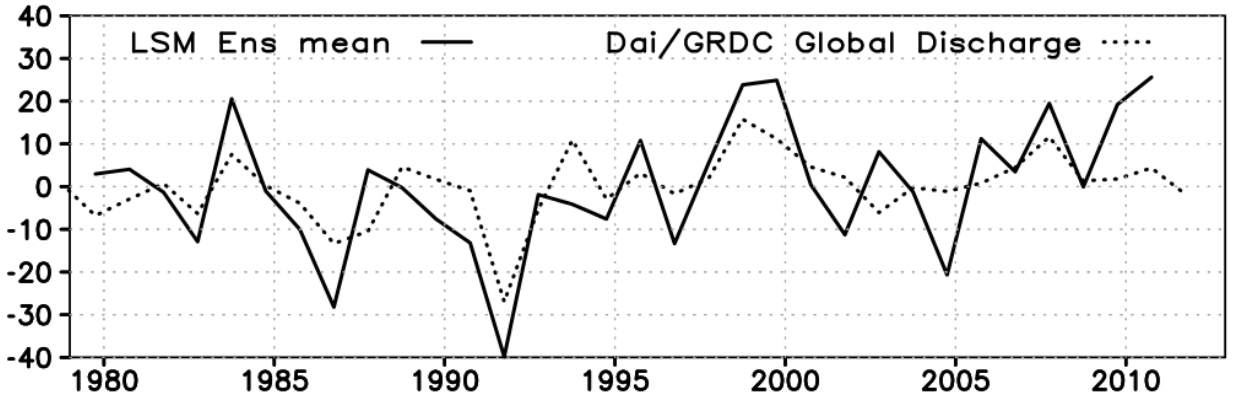


Figure 2. Annual mean anomalies of LSM ensemble mean P-ET and globally-integrated streamflow from Dai (2016). Units are mm yr^{-1}. LSM monthly anomalies from Figure 1 have been summed over water year intervals beginning in October 1979.

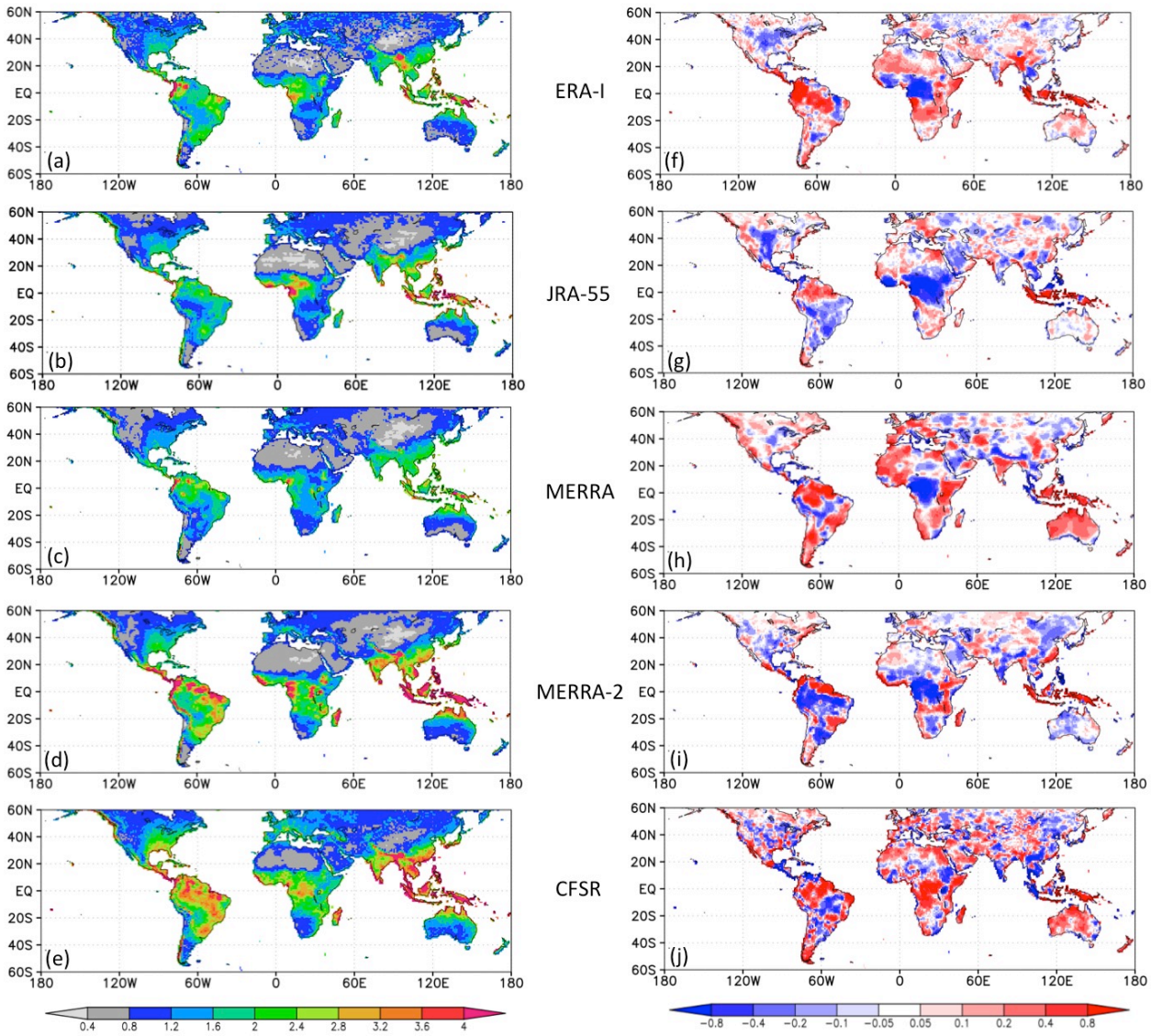


Figure 3. Statistics for monthly mean VMFC anomalies for various reanalyses over the period 1979-2012. *Left (a-e):* RMS of deviations (mmd^{-1}). *Right (f-j):* Trends ($mmd^{-1}decade^{-1}$).

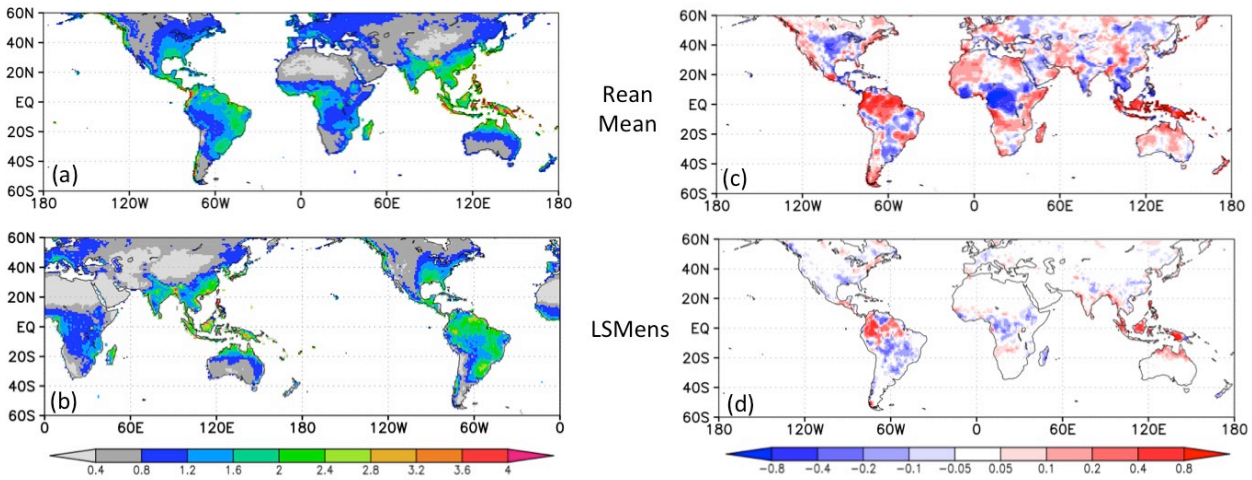
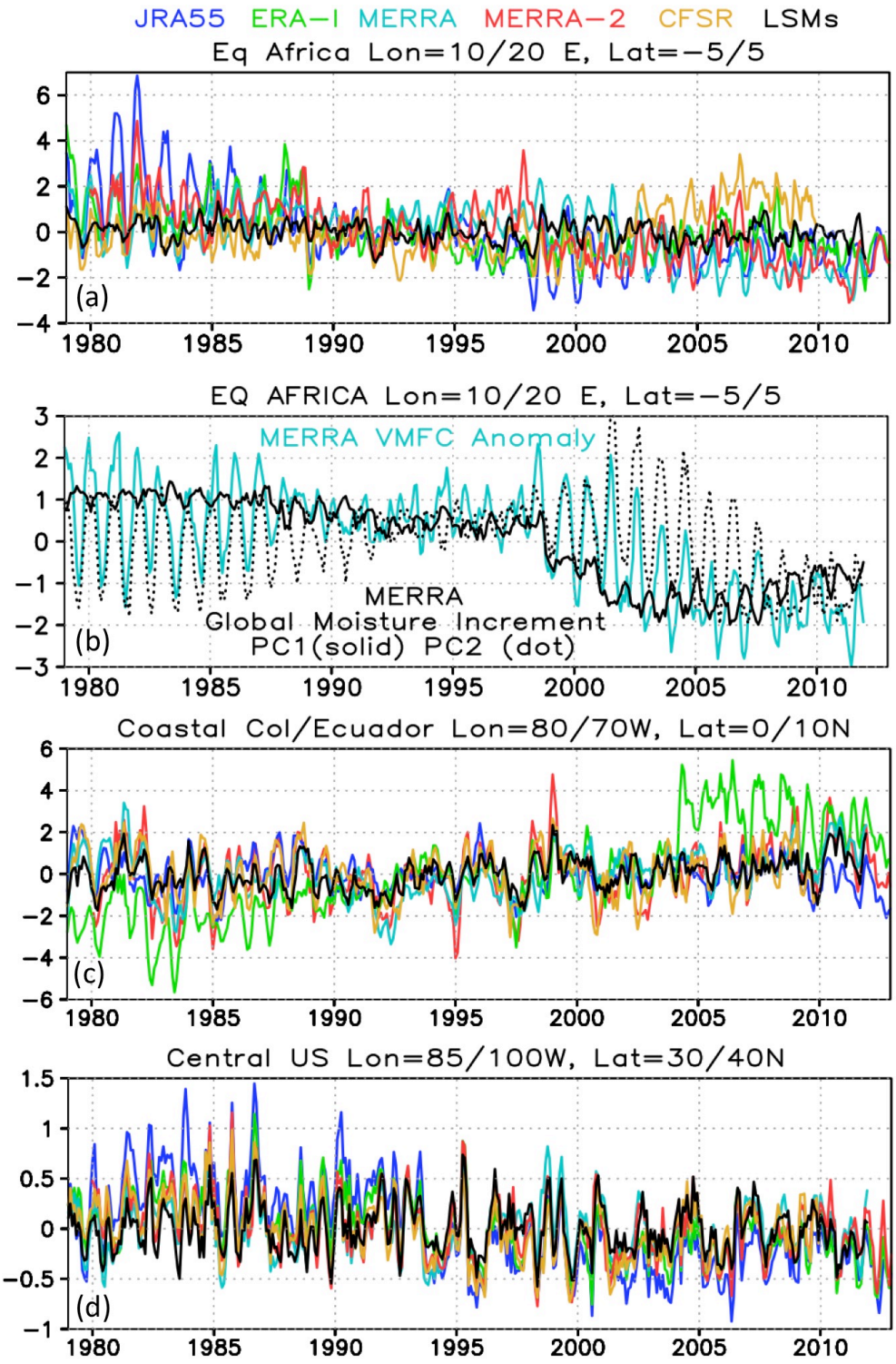


Figure 4. Ensemble mean statistics for reanalysis and LSM monthly anomalies over the period 1979-2012. *Left panel*: RMS (mmd^{-1}) . *Right Panel:* Trends ($mmd^{-1}decade^{-1}$).

1110



1111

1112 Figure 5. Time series of reanalyses VMFC and ensemble LSM anomalies (mm d^{-1}) over (a)
1113 Equatorial Africa (c) Coastal Colombia / Ecuador region, and (d) Central U.S. (b) Shows MERRA
1114 VMFC (red) and first two global PCs of vertically-integrated moisture increment.

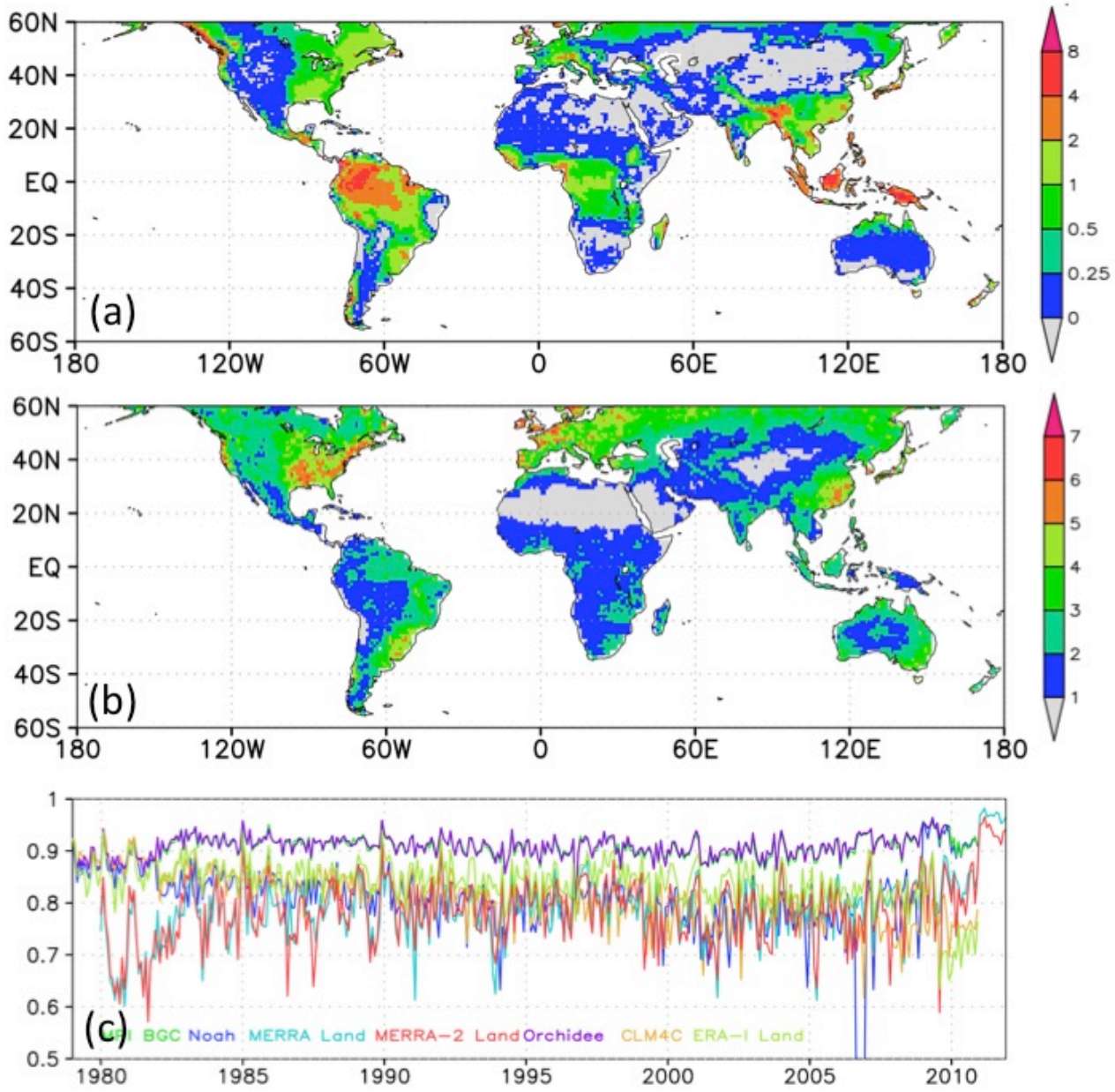


Figure 6. (*a*) Ensemble mean climatological P-ET (mmd^{-1}). (*b*) S / N for LSMs (see text for details on calculation). (*c*) P-ET ACC time series of each LSM with the ensemble mean.

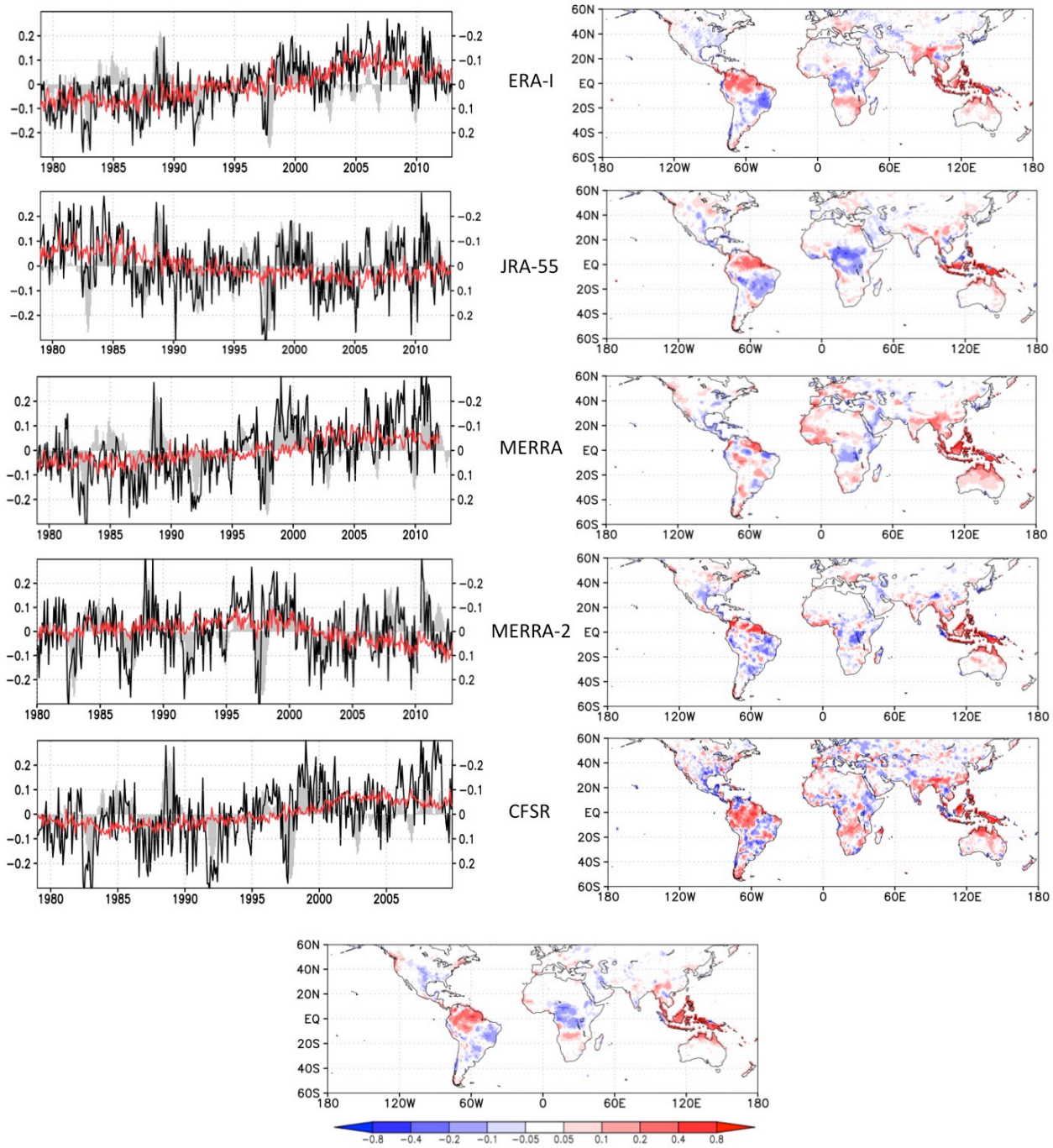


Figure 7. (*Left*) Time series of globally averaged, 60° N/S, reanalysis VMFC (black) and area-averaged corrections (red). Niño 3.4 SST x 0.1 is plotted as gray shading with inverted scale on right hand side. (*Right*) Trends (mmd^{-1} $decade^{-1}$) in VMFC* over the period 1979-2012 for various reanalyses after corrections have been applied. Compare to Figure 3, *right* which are the uncorrected trends. (Bottom) Ensemble mean corrected VMFC* trends.

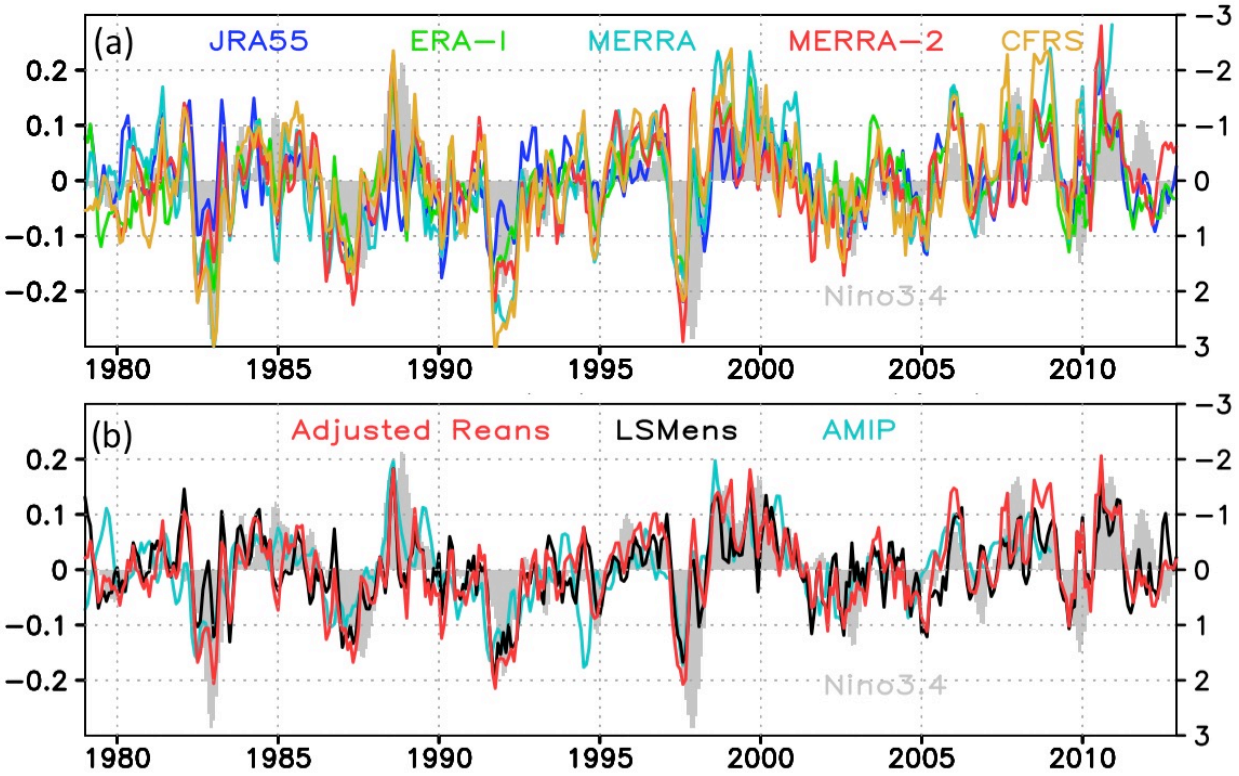


Figure 8. (*a*) Time series of individual corrected reanalysis VMFC* global land area-average (60°N/S). (*b*) P-ET from individual LSMs (black) and mean VMFC* from corrected reanalyses (red) and AMIPs (cyan). A three-month running smoothing is applied. Units are mmd $^{-1}$. Niño 3.4 SST anomalies are plotted in gray shading with inverted scale (deg C) on right.

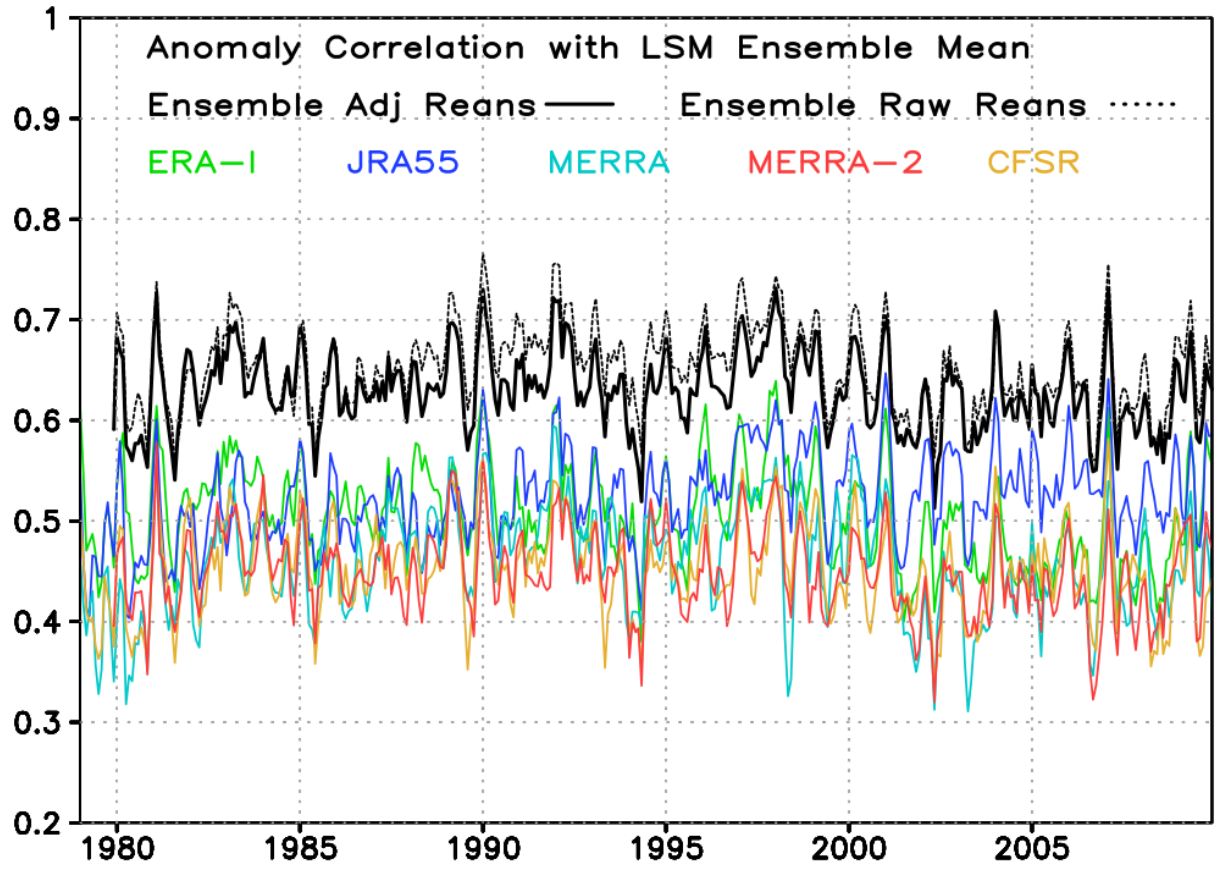


Figure 9. Anomaly correlations, ACC, between individual corrected reanalysis VMFC* and ensemble mean LSM P-ET. ACC for the ensemble corrected and uncorrected reanalyses with the ensemble LSMs is shown by the solid (dotted) black lines, respectively. A running 3-month smoother has been applied to each time series for display.

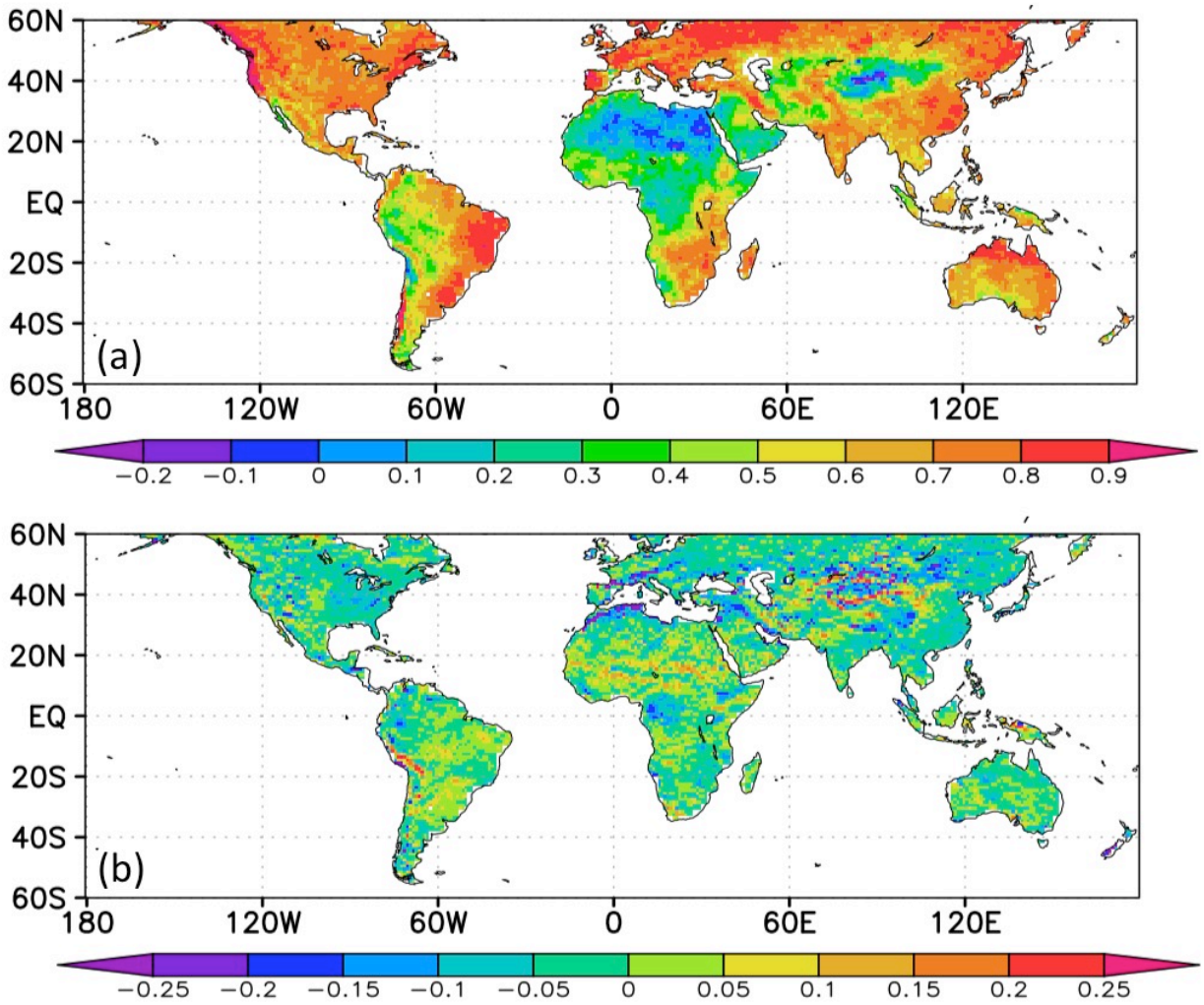


Figure 10. Local correlations (1979-2012): (a) Ensemble mean adjusted reanalysis VMFC* and ensemble mean LSM P-ET, (b) Adjusted VMFC* correlation with LSMs minus raw reanalysis ensemble reanalyses VMFC correlation with ensemble LSM P-ET. Note different color scales.